1982

Attainable bounds for generalized moments via mathematical programming

Michael D. Hale
Iowa State University

8221192

Hale, Michael D.

ATTAINABLE BOUNDS FOR GENERALIZED MOMENTS VIA MATHEMATICAL PROGRAMMING

Iowa State University PH.D. 1982

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

Attainable bounds for generalized moments via mathematical programming

by

Michael D. Hale

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1982

TABLE OF CONTENTS

I. INTRODUCTION

1. Preliminaries

There are numerous problems in decision analysis. Even when one has clearly defined objectives, the choice among alternative actions is usually complicated by uncertainty in the consequences of those actions. Cardinal utility theory is often used to aid in the decision process by the application of probabilistic notions. However, it is frequently the case that utility theory is difficult to apply because some of the probability distributions are unknown or incompletely specified. The present work is the outgrowth of an attempt to clear up some of the uncertainty when one knows nothing of a probability distribution, save its first two moments.

This chapter will briefly cover some of the basic ideas of utility theory. The basic problem will be explained in more detail and some previous attempts at solution will be noted. Additionally, some notions of mathematical programming will be presented.

2. Utility Theory

The basic reason for utility theory is that, generally speaking, a person's strength of preference for payoffs of a gamble is non-linear. This observation came about because mathematical expectation did not work well when payoffs were large sums of money. For example, almost any poor man would prefer one million dollars with certainty to a ten percent chance of receiving ten million dollars.

This phenomenon was apparently well-known in 1738 when Bernoulli (1954) published his famous St. Petersburg paper. In this paper, he proposed a technique to evaluate the "utility" of a lottery. This technique is mathematically equivalent to taking the expectation of $\log_e (c + X)$, where X is the payoff, or consequence, and c is the person's base capital. It should be noted that Bernoulli's method is the generalization of a technique which Cramer outlined in a letter to Bernoulli in 1728, although Bernoulli (1954) apparently formulated his own method independently.

A natural extension of Cramer's and Bernoulli's idea is the fruitful notion of utility function. In a given context, suppose that the set of consequences R is completely ordered by a preference ordering. Then the real valued function U defined on R is a utility function if U is such that for every pair a_1, a_2 of actions with consequences in R, a_1 is not preferred to a_2 if and only if

$$\int_R U(r)dF_1(r) \leq \int_R U(r)dF_2(r) \tag{1.2.1}$$

where F_i is the probability distribution over R when action i is adopted. This proves very useful when evaluating an action with a large number of possible payoffs.

It was not until fully two hundred years after Cramer's and Bernoulli's work that an axiomatic structure for utility functions was introduced. This was done independently by Ramsey (1931) in a somewhat obscure fashion and by von Neumann and Morgenstern (1947) in a highly popular book. Particularly significant in these two expositions is the development of an axiomatic structure which provided the framework for proving the existence of additive utility functions.

If one accepts the von Neumann and Morgenstern axioms, then their results are completely valid. There has been heated discussion over these axioms, however, and several other authors have offered alternate axioms. An excellent comparison of the leading systems may be found in MacCrimmon and Larsson (1975). Even though some insist that 1.2.1 is an incorrect rule for choice (see Hagen (1979)), they are usually dismissed as in Amihud (1974), sometimes curtly (c.f. DeFinetti (1979)). The maximum expected utility school is by far the most popular and that approach, in the von Neumann-Morgenstern formulation, will be followed herein.

Fishburn (1968, 1969) strengthened the applicability of this line by extending the theory to include more general probability distributions over more general sets. (The original von Neumann-Morgenstern structure dealt specifically with denumerable lotteries.) More recently, Keeney and Raiffa (1976) have shown that utility theory is feasible for situations where each consequence is actually an n-tuple whose elements each signify an area of concern to the decision maker. This multi-attribute utility theory has been growing in popularity and has proven useful in situations with a high degree of complexity. An example of this is in the selection of safety programs for nuclear reactor safety as described by Ritzman and Husseiny (1980). Not everyone is enthusiastic about this extension, however, as for example Leung (1978) who questions the worth of the additional complexity required. This present work will not delve into these multi-attribute issues, and will be restricted to what is commonly referred to as univariate utility theory.

To effectively apply utility theory one must know something about the shape of one's utility function, and also the payoff distributions corresponding to the available actions. Much research has concentrated on determining the form of an individual's utility function. The work of Mosteller and Nogee (1951), as well as Davidson, et al. (1957), addressed some of the problems involved and made some pioneering efforts in axiom verification and preference structuring. Becker, et al. (1964) used a sequential technique to assess the form of an individual's utility function, and Biorn (1974) estimated the flexibility of marginal utility of money from aggregate consumption data. Thus, while some problems remain, the problem has been the subject of much research, much of which suggests that utility functions can be ascertained, at least to within reasonable approximation. The sense of the present work will be to assume this, and to address the equally challenging problems that arise when the payoff distributions of interest are not well known.

Two approaches have been taken in dealing with the problem of incomplete knowledge of utility or payoff distribution functions. The first is the stochastic dominance technique. Several results for this method are to be found in Tesfatsion (1976), Hadar and Russell (1969), Whitmore (1970), Borch (1979), Chipman (1973), or Russell and Sev (1978). Two interesting papers dealing with computer implementations of the method are to be found in Porter, et al. (1973) and Bawa, et al. (1979). The reader interested in fundamentals is referred to Fishburn and Vickson (1978) or Brumelle and Vickson (1975). This theory assumes that payoff distributions are well known but that the utility function is known only to the extent of satisfying certain assumptions such as concavity, monotonicity, etc. As indicated above, many people do believe that a utility function can be determined, at least to within good approximation. An interesting question is whether one can actually assume complete knowledge about various payoff distributions if one cannot accurately assess one's own preferences for those payoffs. It is worthwhile to note in passing that the weak duality approach might be employed to solve stochastic dominance problems in a similar manner to that presented in this work.

The other outlook, and by far the more popular, is the mean-variance approach. The utility function is assumed to be known but the payoff distributions are unknown, although usually known to within the first two moments, as the name suggests.

The method was first proposed by Markowitz (1952). He described an E-V efficient set as a certain subset of the boundary of the set of mean-variance pairs for the payoff distributions corresponding to all available actions. See also Baron (1977). This subset has the property that, if a distribution is in the subset then no available distribution with the same mean has smaller variance and no available distribution with the same variance has larger mean. See section three of chapter four for more discussion of this approach. The method is especially appropriate if the investor is constructing a portfolio and can allocate any incremental value to a stock or bond, but is less so if he must choose from only a denumerable set of alternatives. As is pointed out in section three of chapter four, this method is generally good, but is susceptible of misapplication.

The efficient set is designed for a decision maker with a concave non-decreasing utility function. Such a decision maker is said to be risk averse, possessing a utility function $U(r)$ such that for every probability distribution F on R,

$$\int_R U(r)dF(r) \leq \int_R r dF(r) \ . \qquad (1.2.2)$$

In practice, utility functions are usually concave non-decreasing. There are times, however, when a convex utility function is required to represent a gambling attitude or when a low payoff is highly desirable. For convenience most utility functions are scaled to the interval $[0,1]$ or $[0,+\infty)$. Often one takes $U(0) = 0$ and $U(1) = 1$. This is particularly useful for purposes of comparison and for multiattribute utility considerations. It is permissible, of course, since cardinal utility generally is agreed to be determinable only to within a linear transformation. See Arrow (1974) for an alternate viewpoint.

Another aspect of utility theory worth mentioning is the property of decreasing absolute risk aversion (DARA). It is exhibited when the willingness to engage in small bets of fixed size increases with wealth. A necessary (but not sufficient) condition for a utility function to be DARA is for $U'''(x) > 0$ on the domain of x i.e., U' is convex on the domain of x. The theory of chapter two is thus applicable for any DARA utility function on a finite or semi-infinite interval.

A popular variant of E-V analysis involves the approximation of expected utility. This is typically done by some function of U, μ, and σ^2, but might involve more (Samuelson 1970). Loistl (1976) points to certain problems associated with this approach. Much of the theory has been coordinated by Levy and Markowitz (1979) through their approximator f_k. This approximator is examined in the fifth section of chapter three. Its effectiveness is limited in failing to take into account bounds of the sort developed in this thesis. Approximation is convenient, however, and is, at this writing, a virtual necessity for the application of multiattribute utility theory.

To this point almost nothing has been published concerning bounds for $E\{U(X)\}$ when U is known but the payoff distribution is unknown. The work of Mantell (1976) is examined in section four of chapter three. It is the only paper known to the author to specifically address this heretofore surprisingly neglected aspect of the theory of utility.

3. The Motivating Problem

This section presents the specific problem which motivated much of this work. Also presented is a general version of this problem. Chapter two deals with that more general version, reserving treatment of the motivating problem for chapter three, where it is viewed as a special application of the general theory.

Suppose that an investor has a utility function U on the interval [a,b] of possible payoffs. An action is under consideration. The consequences of that action are uncertain, however, so that the payoff has unknown distribution, F . The investor is reasonably certain of the first and second moments of F . Can this knowledge be utilized to gain more information about E{U(X)}?

There are techniques available for approximating E{U(X)}, as are described by Levy and Markowitz (1979). How good is such an approximation, however? Is there any guarantee of "closeness"? Do bounds exist for E{U(X)}? If so, what are they? The present work was initiated in an attempt to answer such questions.

The problem is addressed by considering the following more general problem. Let g be a real-valued function on [a,b] . Then find

$$\min_{F \in \mathcal{F}} \int_a^b g\,dF$$

where $\mathcal{F} = \{F \mid F \text{ is a cdf on } [a,b] \text{ with } \int_a^b x\,dF = \mu_1,\ \int_a^b x^2\,dF = \mu_2\}$. This problem is one of a larger class often referred to as generalized Chebyshev inequalities. Chapter two is devoted to characterizing a class of functions, g, for which a certain weak duality approach yields a solution. There it will be seen that the class is actually rather broad, including many familiar functions such as $\log_e x$, $\exp(x)$, x^y, and y^x .

Another approach is briefly presented in section four of chapter four. It should be noted that the technique has many applications not related to utility theory. See Kim (1978) or Pukelsheim (1978) for some other uses. Also see chapter four of this present work, as well as Kim (1979), for possible reliability applications. Other interesting applications are given by David and Kim (1979) for information functionals and by Kim and David (1979) for large deviations in Markov processes.

The general problem is still unsolved, although Karlin and Studden (1966) showed that mathematical programming may be applied to solve any such problem. In the Karlin and Studden approach, each g is considered on its own merits for solution. Kim (1979) pointed to a method of solution for a certain class of g's. As mentioned earlier, chapter two of this present work characterizes a class of g's for solution. Another class is presented in section four of chapter four.

Actually, it is probably unreasonable to expect more than method-class pairs for solution, where it is understood that "method" includes specification of exactly how an optimal distribution is obtained. The method-class pair described in chapter four is not completely explored. That is, the method will work for some g's not in the given class. The method of chapter two, however, works for all μ_1,μ_2 pairs if and only if g is a member of the pertinent specified class. The background for the approach taken in this thesis is presented in the following section.

4. Weak Duality

Some rudimentary ideas of mathematical programming are presented in this section. It is intended to be merely a very brief introduction, presenting the notion of weak duality in a simple manner. Also included is an indication of how weak duality may be applied to the general problem of the last section.

Consider then a problem

$$P : \underset{x \varepsilon \Lambda}{\text{minimize}}\ f(x) \ .$$

It is useful to consider in conjunction with P a problem

$$D : \underset{y \varepsilon \Omega}{\text{maximize}}\ h(y) \ ,$$

often taken to be the Lagrangian dual of P, such that

$$f(x) \geq h(y) \quad \forall\, x \in \Lambda,\ y \in \Omega \ ,$$

in which case D is said to be weakly dual to P .

The notion of weak duality has proven quite useful as a tool for optimization e.g., see most of the works referenced in the preceding section. In particular, when $x^0 \in \Lambda$, and $y^0 \in \Omega$ are such that $f(x^0) = h(y^0)$, then x^0 is an optimal solution for P . David and Kim (1979), for example, systematically exploit this fact for a variety of optimization problems. This same fact is utilized in similar fashion in chapter two. In particular, P and D become

$$P_1 : \min_{F \varepsilon \mathcal{F}} \int_a^b g\,dF$$

where $\mathcal{F} = \{F | F \text{ is a cdf on } [a,b] \text{ with } \int_a^b x dF = \mu_1, \int_a^b x^2 dF = \mu_2\}$,

and

$$D_1: \max_{(\lambda_0,\lambda_1,\lambda_2)\varepsilon L} (\lambda_0+\lambda_1\mu_1+\lambda_2\mu_2),$$

where $L = \{(\lambda_0,\lambda_1,\lambda_2) | (\lambda_0,\lambda_1,\lambda_2) \ \varepsilon\ E^3,\ Q(x) = \lambda_0 + \lambda_1 x + \lambda_2 x^2 \leq g(x)$

$\forall\ x \varepsilon [a,b]\}$.

A tentative solution F^0, $(\lambda_0^0,\lambda_1^0,\lambda_2^0)$ is presented so that $\int_a^b g dF^0 = \lambda_0^0 + \lambda_1^0\mu_1 + \lambda_2^0\mu_2$, and most of the work of this thesis centers on finding conditions on g insuring that $(\lambda_0^0,\lambda_1^0,\lambda_2^0)$ is feasible, i.e., is an element of L .

The reader may find greater depth on programming in Rockafellar (1974) and Sposito (1975). See also Eggleston (1958). See also Marshall and Olkin (1960a, 1960b) for additional material on generalized Chebyshev inequalities, as well as Karlin and Studden (1966), and particularly Mallows (1956). Related work may be found also in Kingman (1963), Isii (1960, 1963, 1964), and Rogosinski (1958). The field is quite broad and many researchers have contributed to it. The papers referenced here contain numerous further references for the interested reader.

5. Overview

The purpose of this work is to solve the problems stated in section three of the present chapter. The approach taken was outlined in section four. Chapter two examines a solution in detail, presenting several conditions which guarantee that the approach works. Some of the conditions may appear rather difficult to verify. Others are quite simple, such as convexity of g', for example.

Chapter three relates the results of chapter two to utility theory. There it is seen that utility functions usually are members of the class of functions for which the proposed method works, and current approximation theory is examined in light of the bounds obtainable thereby.

Finally, chapter four contains numerous examples and extensions. The purpose of that chapter is to clarify many of the previously presented ideas as well as suggesting ideas for further consideration. In particular, examples are taken from reliability and utility. The general problem is briefly examined again and another method of solution discussed.

II. THE GENERAL FORMULATION

1. Introduction

In this chapter, the general problem of finding the extreme of

$$\int_a^b g dF$$

is considered, where g is a known function and F is a distribution function with known first and second moments. A weak duality approach to this problem is presented, and necessary and sufficient conditions for g are presented in order for certain bounds derivable by this approach to be attainable. The result of this is the characterization of the class of functions for which this weak duality approach will yield attainable bounds for arbitrary first and second moments.

The usefulness of this characterization is readily apparent; if a function is a member of this class, then intuitive guesses, verifying trial solutions, etc. are unnecessary. One only needs to apply the method developed in this chapter to obtain the bounds as functions of the first two moments of F.

The following section presents the weak duality approach to be used.

2. The Weak Duality Framework

Suppose, for given g, that one has the problem

$$P_1: \quad \underset{F}{\text{minimize}} \int_a^b g dF \tag{2.2.1}$$

$$\text{subject to} \quad \int_a^b dF = 1 \tag{2.2.2}$$

$$\int_a^b x\, dF = \mu_1 \tag{2.2.3}$$

$$\int_a^b x^2 dF = \mu_2 \tag{2.2.4}$$

$$F(-\infty) = 0,\ F \text{ right continuous} \tag{2.2.5}$$

$$dF \geq 0 \tag{2.2.6}$$

Conditions 2.2.2, 2.2.5, and 2.2.6 insure that F is a cdf. In 2.2.3, 2.2.4, μ_1 and μ_2 are given finite constants.

It is easily shown that the following problem is weakly dual to P_1.

$$D_1: \quad \underset{(\lambda_0,\lambda_1,\lambda_2)}{\text{maximize}} \quad \lambda_0 + \lambda_1\mu_1 + \lambda_2\mu_2 \tag{2.2.7}$$

subject to

$$Q(x) \equiv \lambda_0 + \lambda_1 x + \lambda_2 x^2 \leq g(x) \ \forall\, x \in [a,b] \tag{2.2.8}$$

$$(\lambda_0,\lambda_1,\lambda_2) \in E^3 \tag{2.2.9}$$

This is done as follows. Suppose that F is feasible for P_1 and $(\lambda_0,\lambda_1,\lambda_2)$ is feasible for D_1. That is, F satisfies 2.2.2-2.2.6 and $(\lambda_0,\lambda_1,\lambda_2)$ satisfies 2.2.8, 2.2.9. Then by 2.2.8

$$\int_a^b g\,dF \geq \int_a^b Q\,dF \tag{2.2.10}$$

$$= \lambda_0 + \lambda_1\mu_1 + \lambda_2\mu_2 \tag{2.2.11}$$

since F satisfies 2.2.2-2.2.4,

so that D_1 is weakly dual to P_1. That is, for F feasible for

P_1 and $(\lambda_0,\lambda_1,\lambda_2)$ feasible for D_1, the objective function of P_1 is greater than or equal to the objective function of D_1.

Hence,

$$\min_{F \text{ feas.} P_1} \int_a^b g\,dF \geq \max_{\substack{(\lambda_0,\lambda_1,\lambda_2) \\ \text{feas.} D_1}} (\lambda_0+\lambda_1\mu_1+\lambda_2\mu_2) \,. \tag{2.2.12}$$

Of course, as noted in Chapter I, if there exists an F^* feasible for P_1 and a $(\lambda_0^*,\lambda_1^*,\lambda_2^*)$ feasible for D_1 such that

$$\int_a^b g\,dF^* = \lambda_0^* + \lambda_1^*\mu_1 + \lambda_2^*\mu_2 \tag{2.2.13}$$

then F^* is optimal for P_1 and $(\lambda_0^*,\lambda_1^*,\lambda_2^*)$ is optimal for D_1.

As several authors (c.f. David and Kim (1979), Pukelsheim (1978), and Kim (1978)) have noted, if $(\lambda_0,\lambda_1,\lambda_2)$ is feasible for D_1 and F is feasible for P_1 with all of its points of increase at points where $Q = g$, the so called "osculating set", then 2.2.13 is satisfied.

That is,

$$\int_a^b (g-Q)\,dF = 0 \tag{2.2.14}$$

since F has increase only when the integrand is zero. Thus, a sufficient condition for feasible F and feasible $(\lambda_0,\lambda_1,\lambda_2)$ to be be optimal is that F has all of its mass on points where $Q = g$.

Naturally the difficulty lies in finding such an F and $(\lambda_0,\lambda_1,\lambda_2)$.

In this chapter, a method is developed to construct a feasible F and $(\lambda_0,\lambda_1,\lambda_2)$ satisfying 2.2.14, contingent, of course, on g

satisfying certain conditions to be presented later.

The next section proposes an F and a $(\lambda_0,\lambda_1,\lambda_2)$ which satisfy 2.2.14. By construction F will always be feasible, but the feasibility of $(\lambda_0,\lambda_1,\lambda_2)$ depends on the "shape" of g. Conditions for feasibility are presented in Section 4.

It should be noted that in much of the ensuing discussion the terminology "feasible Q" will be used to signify "feasible $(\lambda_0,\lambda_1,\lambda_2)$".

3. The Proposed Extremal cdf and Q

With a view to the minimization 2.2.1, it is natural to search for an extremal cdf.

Define

$$\mathfrak{F} = \{F \mid F \text{ is a cdf on } [a,b] \text{ with finite first } (\mu_1), \text{ and second } (\mu_2) \text{ moments}\}. \tag{2.3.1}$$

Also define

$$\mathfrak{F}_c = \{F \mid F \in \mathfrak{F}, \text{ and } F \text{ is a one-point distribution, or a two-point distribution with mass at } c\}. \tag{2.3.2}$$

Also let $$\bar{\mathfrak{F}} = \{\mathfrak{F}_c \mid c \in [a,b]\}. \tag{2.3.3}$$

At this time, a few things should be noted. It is assumed that a is finite, but b may be $+\infty$, in which case $[a,b]$ should of course be written $[a,+\infty)$. Also, by definition of $\mathfrak{F}_c$, $c \in [a,b]$.

The following lemma shows why $\mathfrak{F}_a$ (and $\mathfrak{F}_b$ if b is finite) is useful.

Lemma 2.3.1

$\mathfrak{F}_a$ is the only member (except for $\mathfrak{F}_b$, if b is finite) of $\bar{\mathfrak{F}}$ that has for every $F \varepsilon \mathfrak{F}$, an element F^0 such that

$$\int_a^b xdF = \int_a^b xdF^0 \tag{2.3.4}$$

and

$$\int_a^b x^2 dF = \int_a^b x^2 dF^0. \tag{2.3.5}$$

Proof: (By contradiction)

Suppose that for some $c \varepsilon (a,b)$ the class $\mathfrak{F}_c$ has the property that for every $F \varepsilon \mathfrak{F}$ there exists $F^0 \varepsilon \mathfrak{F}_c$ with conditions 2.3.4 and 2.3.5 satisfied. Then choose $F \varepsilon \mathfrak{F}$, $\mu_2 \varepsilon (c^2, b^2]$ such that

$$\int_a^b xdF = c \tag{2.3.6}$$

and

$$\int_a^b x^2 dF = \mu_2 \tag{2.3.7}$$

By definition of $\mathfrak{F}$, there must exist F satisfying 2.3.6, 2.3.7 for $c \varepsilon (a,b)$. But for $F^0 \varepsilon \mathfrak{F}_c$, $\int_a^b xdF^0 = c(1-p) + zp$ for some $p \varepsilon [0,1]$, $z \varepsilon [a,b]$. So 2.3.6 $\Rightarrow$ $p = 0$ or $z = c$. Hence, $\int_a^b x^2 dF^0 = c^2(1-p) + z^2 p = c^2 < \mu_2$, and condition 2.3.5 is therefore not satisfied. Thus, $\mathfrak{F}_c$ cannot possess the above property and a contradiction results.

Hence, no $\mathfrak{F}_c$, $c \varepsilon (a,b)$, contains an F^0 for every $F \varepsilon \mathfrak{F}$ with F^0 satisfying 2.3.4 and 2.3.5.

It now remains to show that $\mathfrak{F}_a$ does have this property.

For $F \in \mathfrak{F}$ with moments μ_1, μ_2 it is necessary to find p, z such that

$$a(1-p) + zp = \mu_1 \qquad (2.3.8)$$

and

$$a^2(1-p) + z^2 p = \mu_2 \qquad (2.3.9)$$

If $\mu_2 = \mu_1^2$ then $p = 1$ and $z = \mu_1$. Otherwise, solving 2.3.8 and 2.3.9 for z, p yields

$$z = (\mu_2 - a\mu_1)/(\mu_1 - a) \qquad (2.3.10)$$

$$p = (\mu_1 - a)^2/(\mu_2 - 2a\mu_1 + a^2) \qquad (2.3.11)$$

Notice that 2.3.11 may be rewritten as

$$p = \frac{(\mu_1 - a)^2}{(\mu_1 - a)^2 + (\mu_2 - \mu_1^2)} \qquad (2.3.12)$$

In this rather insightful form, it is clear that $p \in [0,1]$ since $\mu_2 \geq \mu_1^2$. It also follows that $z \geq a$ since

$$z = \frac{(\mu_2 - a\mu_1)}{(\mu_1 - a)} \geq \frac{(\mu_1^2 - a\mu_1)}{(\mu_1 - a)} = \mu_1 \qquad (2.3.13)$$

and $\mu_1 \in [a,b]$.

Finally, is $z \leq b$? Of course, if $b = +\infty$. To show that $z \leq b$ if $b=+\infty$ it is necessary to recall that $F \in \mathfrak{F}$ and that for a given μ_1, $\int_a^b x^2 dG$ is maximized by the distribution with mass at a and b only. That is, the problem

$$\max_{G \text{ a cdf}} \int_a^b x^2 dG \qquad (2.3.14)$$

$$\ni \int_a^b x\,dG = \mu_1 \qquad (2.3.15)$$

is solved by G^0 where G^0 puts mass $(\mu_1-a)/(b-a)$ at b and mass $(b-\mu_1)/(b-a)$ at a. This is shown below since the right hand side of the equality $\int_a^b x^2 dG^0 = (b+a)\mu_1 - ab$ is a constant. (2.3.16)

Subtracting this from the objective function one gets the equivalent problem

$$\max_{G \text{ a cdf}} \int_a^b (x^2-(a+b)x + ab)\, dG \qquad (2.3.17)$$

$$\ni \int_a^b x dG = \mu_1$$

and 2.3.17 may be rewritten as

$$\max_{G \text{ a cdf}} \int_a^b (x-a)(x-b)\, dG\,. \qquad (2.3.18)$$

Since $x \in [a,b]$, the integrand has a maximum of zero, which is achieved only at $x = a$ or $x = b$, so that indeed G^0 solves the problem as was claimed. Therefore, $z \leq b$ and $\mathfrak{F}_a$ performs as stated.

Note that, for b finite, $\mathfrak{F}_b$ performs in the same manner as $\mathfrak{F}_a$. This is useful since, for b finite, a member of $\mathfrak{F}_a$ will be used to obtain one bound, and a member of $\mathfrak{F}_b$ will be used to obtain a bound from the other side, i.e., one will be an upper, the other a lower bound.

As was stated at the end of section two, the aim is to construct a feasible F and a feasible $(\lambda_0,\lambda_1,\lambda_2)$. The member of $\mathfrak{F}_a$ satisfying 2.2.3 and 2.2.4 will be the feasible F^* proposed as an optimal cdf. The aim of the following discussion is to develop a feasible $(\lambda_0^*,\lambda_1^*,\lambda_2^*)$ corresponding to this F^* which will satisfy 2.2.14.

fol

Lem

ext

Pro

and

con

sa

and

si

pos

fea

Thus, the quadratic Q^* will coincide with g at the points a and z. When, as is normally the case, z is an interior point of $[a,b]$, Q^* will be tangent to g at z.

These ideas are proven through assumptions concerning g.

Condition 2.3.1:

The function g is absolutely continuous on $[a,b]$ and has continuous derivative on (a,b).

This means that g may be expressed as

$$g(x) = g(a) + \int_a^x g'(y)dy, \quad x \in [a,b) \,. \tag{2.3.19}$$

The following definition will be used extensively:

Definition 2.3.1

For each $t \in (a,b)$

define

$$Q_t(x) = \lambda_0 + \lambda_1 x + \lambda_2 x^2 \tag{2.3.20}$$

such that $$Q_t(a) = g(a) \tag{2.3.21}$$

$$Q_t(t) = g(t) \tag{2.3.22}$$

$$Q_t'(t) = g'(t) \tag{2.3.23}$$

Of course this might be denoted as $Q(x;a,t,g)$ more properly, but for simplicity will simply be written as Q_t or $Q_t(x)$ since g and a are considered to be fixed.

That one can always construct a unique Q_t is proven in the following lemma.

Lemma 2.3.2

Let g satisfy Condition 2.3.1. Then for every $t \varepsilon (a,b)$ there exists a unique quadratic function Q_t satisfying 2.3.21-2.3.23.

Proof:

Conditions 2.3.21-2.3.23 may be rewritten in matrix form as

$$\begin{bmatrix} 1 & a & a^2 \\ 1 & t & t^2 \\ 0 & 1 & 2t \end{bmatrix} \begin{bmatrix} \lambda_0 \\ \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} g(a) \\ g(t) \\ g'(t) \end{bmatrix} \qquad (2.3.24)$$

The matrix on the left is non-singular, having determinant $-(a-t)^2$, and the vector on the right has finite components since g satisfies condition 2.3.1. Therefore, there exists a unique vector $(\lambda_0,\lambda_1,\lambda_2)$ satisfying 2.3.24 and Q_t is uniquely determined.

The way is clear now. $F^* \varepsilon \mathfrak{F}_a$ is chosen so that it has first and second moments μ_1,μ_2 respectively. Then Q_z will satisfy

$$\int_a^b (g-Q_z)dF^* = 0$$

since F^* has mass at a and z only, and $g-Q_z$ is zero at those points.

The remaining question, of course, is whether or not Q_z is feasible, i.e., $\leq g$ on $[a,b]$. If so, then Q^* is taken to be Q_z.

Necessary and sufficient conditions for feasibility are detailed in the next section.

If Q_z is feasible, and hence, is the proposed Q^*, then since 2.2.14 is satisfied one has the result that F^* and the corresponding $(\lambda_0^* \; \lambda_1^* \; \lambda_2^*)$ are optimal. Hence, a sharp lower bound for 2.2.1 is given by

$$\int_a^b g\,dF^* = g(a)(1-p) + g(z)p \tag{2.3.25}$$

$$= g(a)\left(\frac{\mu_2-\mu_1^2}{\mu_2-2a\mu_1+a^2}\right) + g(z)\left(\frac{(\mu_1-a)^2}{\mu_2-2a\mu_1+a^2}\right) \tag{2.3.26}$$

or
$$= \frac{g(a)(\mu_2-\mu_1^2) + g\left(\frac{(\mu_2-a\mu_1)}{(\mu_1-a)}\right)(\mu_1-a)^2}{(\mu_1-a)^2 + (\mu_2-\mu_1^2)}$$

by substituting the values for p and z from 2.3.10, 2.3.11. Further details and examples of application will be deferred to a later section since conditions for feasibility have not yet been presented.

One should note here, perhaps, the consequences of restricting oneself to $\mathfrak{F}_a$ and Q_z. If Q_z is feasible, the attainability of the bound implies that one can't do any better on $\mathfrak{F}$ or over all possible quadratics. If Q_z is not feasible, then other, possibly ad hoc, procedures must be invoked.

4. The Fundamental Result

As has been previously stated, in order to apply this weak duality approach it is necessary to find a feasible $(\lambda_0,\lambda_1,\lambda_2)$ and a feasible F such that

$$\int_a^b (g-Q)\,dF = 0 \tag{2.4.1}$$

where

$$Q(x) = \lambda_0 + \lambda_1 x + \lambda_2 x^2,\ x \in [a,b] . \tag{2.4.2}$$

Presented in this section are necessary and sufficient conditions on g for feasibility of Q_z, i.e., for Q_z to satisfy $g \geq Q_z$ on $[a,b]$, (i.e., $g(x) \geq Q_z(x)\ \forall\ x \in [a,b]$) or $g \leq Q_z$ on $[a,b]$. The former is used for an attainable lower bound and the latter for an attainable upper bound, of course.

This section will concentrate on $g \geq Q_z$. The case $g \leq Q_z$ may be treated in parallel fashion, and results concerning this case will be stated without proof.

It is expeditious at this point to enumerate certain facts and relations that either are generally true or are due to the assumptions, and that are used extensively in many of the subsequent proofs. The facts are as follow:

S.1 Q_z' is linear.

S.2 $g' - Q_z'$ is continuous on (a,b).

S.3 If two non-parallel lines cross at the point w, then the line with greater slope lies above the line with lesser slope to the right of w. To the left of w, the line with lesser slope dominates.

S.4 If f and h are continuous functions with $f > h$ at the point w, then there is an open interval containing w on which $f > h$.

The relations, all of which are simple consequences of the constraints 2.3.21 - 2.3.23 which define Q_t are these:

$$\int_a^x g' - Q_t' = g(x) - Q_t(x) \tag{2.4.3}$$

$$\int_a^t g' - Q_t' = 0 \tag{2.4.4}$$

$$\int_a^x g' - Q_t' = \int_x^t Q_t' - g' \ \forall\, x \in (a,t) \tag{2.4.5}$$

$$\int_a^x g' - Q_t' = \int_t^x g' - Q_t' \ \forall\, x \in (t,b] \tag{2.4.6}$$

Sufficient background has now been given for the presentation of the primary theorem of this chapter.

Theorem 2.4.1

Let g be a function satisfying Condition 2.3.1.

Then

$$g \geq Q_t \text{ on } [a,b] \ \forall\, t \in (a,b) \tag{2.4.7}$$

if and only if

$$g' \geq Q_t' \text{ on } [t,b) \ \forall\, t \in (a,b)\ . \tag{2.4.8}$$

A few comments are in order before a proof of this theorem is given. Admittedly there are many functions g not satisfying 2.4.8 for which optimal solutions for P_1 and D_1 may be obtained. One simple example would be the cosine function on $[0,2\pi]$ (see example 4.3.1). In many such situations a solution must be obtained for some specific μ_1,μ_2 pair rather than making a general statement such as 2.3.27, and the problem of finding a feasible pair F, $(\lambda_0,\lambda_1,\lambda_2)$ satisfying 2.4.1 is, in general, not an easy task. This is further complicated if one is trying to construct a bound as a function of μ_1 and μ_2 since, for certain μ_1,μ_2 combinations, a three point extremal distribution may be required rather than a two point distribution (see example 3.1 of Chapter 12, Karlin and Studden (1966)).

If g satisfies 2.4.8, however, then one need not bother with trying to "tailor-fit" a quadratic. For feasibility is assured and 2.3.27 may be taken as the required attainable bound, regardless of the value of μ_1 and μ_2. Thus, the reason for the specification $\forall\ t \in (a,b)$ in the theorem, since t is determined as $(\mu_2-a\mu_1)/(\mu_1-a)$ by the choice of μ_1 and μ_2 (see 2.3.10) and may therefore be any point in the interval (a,b). (More will be said about the cases $t = a$ and $t = b$ at the end of this section.) Put quite simply, if g satisfies 2.4.8 then one need not check for feasibility of Q_t for the particular t (i.e., pair (μ_1,μ_2)) at hand.

Now that motivation has been provided, the proof will proceed.

Proof:

(Necessity) It will be shown that 2.4.7 is incompatible with the existence of points x^0,t^0 where $t^0 \in (a,b)$; and

$$x^0 \in [t^0,b) \tag{2.4.9}$$

such that

$$g'(x^0) < Q'_{t^0}(x^0) \ . \tag{2.4.10}$$

This will be done by showing that the latter condition leads to the existence of a point $\theta \in (a,b)$ such that Q_θ violates 2.4.7.)

Now suppose that 2.4.7 holds (i.e., $g \geq Q_t$ on $[a,b]$ $\forall$ $t \in (a,b)$) and that

for some $t^0 \in (a,b)$ $\exists$ $x^0 \in (t^0,b) \ni g'(x^0) < Q'_{t^0}(x^0)$. (2.4.11)

Now define

$$\theta = \sup_x \{x|g'(x) \geq Q'_{t^0}(x),\ x \in [t^0,x^0)\}. \qquad (2.4.12)$$

Notice that the set is non-empty since, by 2.3.23, it contains the point t^0.

Figure 2.4.1, showing a typical g, may aid in visualizing relations here.

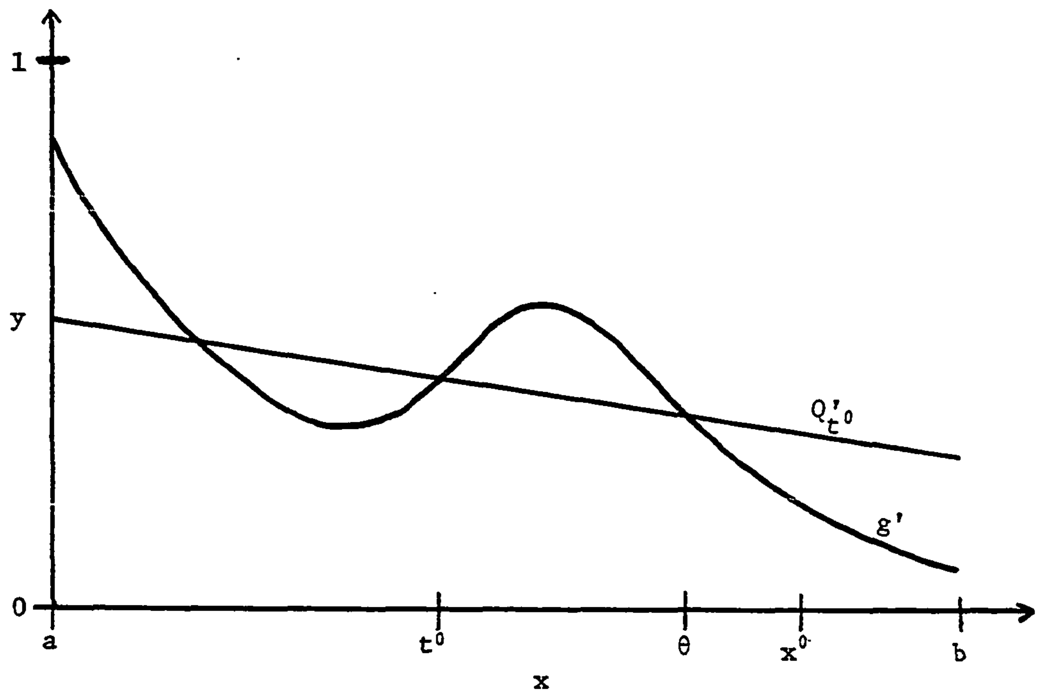


Figure 2.4.1. One possibility for g'

By S.4 there is an open interval containing x^0 on which $Q'_{t^0} > g'$, so that, from 2.4.12,

$$\theta \in [t^0, x^0) \ , \tag{2.4.13}$$

$$g'(\theta) = Q'_{t^0}(\theta) \ , \tag{2.4.14}$$

and

$$g' < Q'_{t^0} \text{ on } (\theta, x^0] \ . \tag{2.4.15}$$

Of course 2.4.15 implies

$$\int_{\theta}^{x^0} g' - Q'_{t^0} < 0 \tag{2.4.16}$$

and, since 2.4.7 holds,

$$\int_{t^0}^{x^0} g' - Q'_{t^0} \geq 0 \tag{2.4.17}$$

by 2.4.6 and 2.4.3. Now

$$\theta > t^0 \tag{2.4.18}$$

by 2.4.13, 2.4.16, and 2.4.17,

and

$$\int_{t^0}^{\theta} g' - Q'_{t^0} + \int_{\theta}^{x^0} g' - Q'_{t^0} = \int_{t^0}^{x^0} g' - Q'_{t^0} \geq 0 \tag{2.4.19}$$

implies, in view of 2.4.16, that

$$\int_{t^0}^{\theta} g' - Q'_{t^0} > 0 \ . \tag{2.4.20}$$

By 2.4.6, this last is equivalent to

$$\int_{a}^{\theta} g' - Q'_{t^0} > 0 \ , \tag{2.4.21}$$

which means that Q_θ will be infeasible violating 2.4.7, as will now be shown.

Consider Q_θ. Statements 2.4.4 and 2.4.21 imply that

$$\int_a^\theta g' - Q'_{t^0} + \int_a^\theta Q'_\theta - g' > 0 \tag{2.4.22}$$

so that

$$\int_a^\theta Q'_\theta - Q'_{t^0} > 0 \quad . \tag{2.4.23}$$

Since Q'_θ and Q'_{t^0} are non-identical straight lines with intersection at θ (see S.1, 2.4.14, 2.3.23), it follows from S.4 that

$$Q'_\theta > Q'_{t^0} \quad \text{on } [a,\theta), \tag{2.4.24}$$

and in particular on the interval $[a,t^0]$. Therefore, by 2.4.24,

$$\int_a^{t^0} Q'_\theta - Q'_{t^0} > 0 \tag{2.4.25}$$

and hence, from 2.4.4,

$$\int_a^{t^0} Q'_\theta - Q'_{t^0} + \int_a^{t^0} Q'_{t^0} - g' > 0 \ . \tag{2.4.26}$$

The left hand side of 2.4.26 may be reexpressed to give

$$\int_a^{t^0} Q'_\theta - g' > 0 \tag{2.4.27}$$

or

$$Q_\theta(t^0) > g(t^0) \tag{2.4.28}$$

which contradicts 2.4.7. Since 2.4.7 and 2.4.11 are incompatible, 2.4.7 implies the complement of 2.4.11, which is precisely 2.4.8.

(Sufficiency)

Suppose that 2.4.8 holds. That is,

$$g' \geq Q'_t \quad \text{on } [t,b) \ \forall\ t \in (a,b)$$

Then for $x \varepsilon [t,b]$

$$\int_t^x g' - Q_t' \geq 0 \tag{2.4.29}$$

so that, indeed,

$$g \geq Q_t \quad \forall\, x \varepsilon [t,b] \quad \forall\, t \varepsilon (a,b).$$

What of the interval (a,t) ?

Again the proof will be by contradiction.

The approach will be to show that if, for some $t \varepsilon (a,b)$, $g < Q_t$ at some point in (a,t) then there is an α such that $g' < Q_\alpha'$ at some point in $(\alpha,b]$.

With this goal in mind suppose that Condition 2.4.8 holds, i.e.,

$$g' \geq Q_t' \text{ on } [t,b) \quad \forall\, t \varepsilon (a,b)$$

and also that

$$\text{for some } z^0 \varepsilon (a,b) \;\exists\; y^0 \varepsilon (a,z^0) \ni g(y^0) < Q_{z^0}(y^0) \;. \tag{2.4.30}$$

Note that 2.4.30 contains the condition

$$\int_a^{y^0} g' - Q_{z^0}' < 0 \;. \tag{2.4.31}$$

Now, at y^0 either $g'(y^0) < Q_{z^0}'(y^0)$ or $g'(y^0) \geq Q_{z^0}'(y^0)$.

The two cases will be considered separately. Suppose

$$g'(y^0) < Q_{z^0}'(y^0) \;.$$

Let

$$\alpha = \inf_x \{x \,|\, g'(x) \geq Q_{z^0}'(x),\; x \varepsilon [y^0,z^0]\}. \tag{2.4.32}$$

The set is non-empty, having z^0 as a member.

One also has

$$\alpha > y^0 \tag{2.4.33}$$

since by S.4 there is an open interval about y^0 on which $g' < Q'_{z^0}$. Continuity yields

$$g'(\alpha) = Q'_{z^0}(\alpha) \tag{2.4.34}$$

and

$$g' < Q'_{z^0} \quad \text{on } [y^0,\alpha) \tag{2.4.35}$$

by definition of α.

So from 2.4.31 and 2.4.35

$$\int_a^{y^0} g' - Q'_{z^0} + \int_{y^0}^{\alpha} g' - Q'_{z^0} < 0 \quad . \tag{2.4.36}$$

This, of course, in conjunction with 2.4.4, 2.4.32, and 2.4.33, implies that

$$\alpha \in (y^0,z^0) \tag{2.4.37}$$

and from 2.4.36 it follows that

$$\int_a^{\alpha} g' - Q'_{z^0} < 0 \ . \tag{2.4.38}$$

This last inequality is the key to the proof of the present result. The proof hinges, actually, on the fact that either $g'(y^0) < Q'_{z^0}(y^0)$ or $g'(y^0) \geq Q'_{z^0}(y^0)$ yields the inequality given in 2.4.38. It will now be shown that the other case, $g'(y^0) \geq Q'_{z^0}(y^0)$, also leads to 2.4.38 for some $\alpha < z^0$.

Figure 2.4.2 illustrates the situation

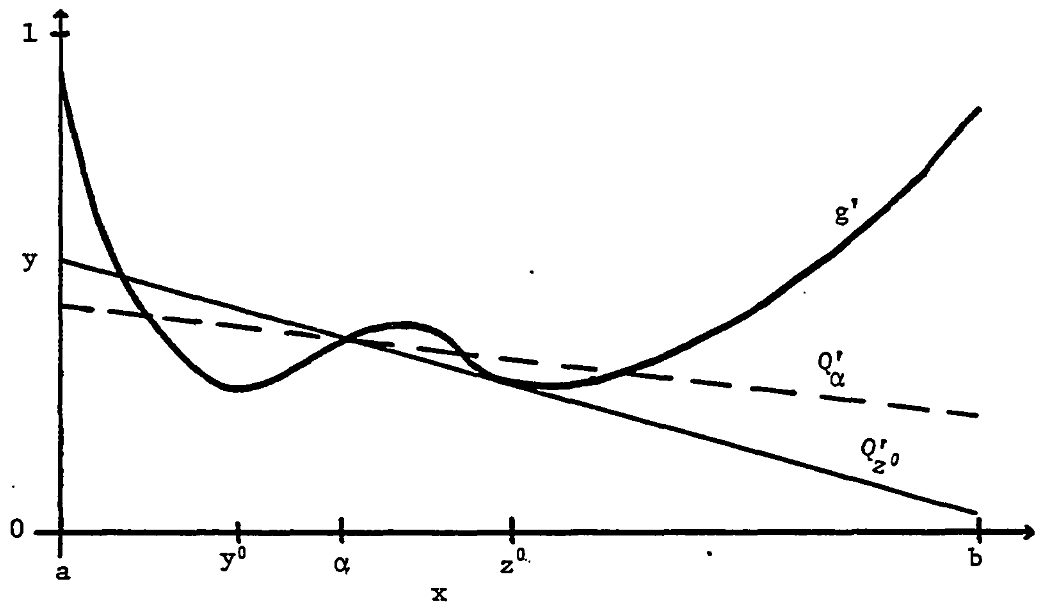


Figure 2.4.2. Another possibility for g'

If $g'(y^0) \geq Q'_{z^0}(y^0)$ let

$$\alpha = \sup_x \{x | g'(x) < Q'_{z^0}(x),\ x \in [a, y^0]\}. \tag{2.4.39}$$

The set is non-empty by 2.4.31, and again by construction

$$g'(\alpha) = Q'_{z^0}(\alpha) \tag{2.4.40}$$

and

$$g' \geq Q'_{z^0} \text{ on } [\alpha, y^0]. \tag{2.4.41}$$

It follows from 2.4.41 that

$$\int_\alpha^{y^0} g' - Q'_{z^0} \geq 0 \tag{2.4.42}$$

and this, in conjunction with 2.4.31, implies

$$\int_a^\alpha g' - Q'_{z^0} < 0\ . \tag{2.4.43}$$

Figure 2.4.3 illustrates this situation.

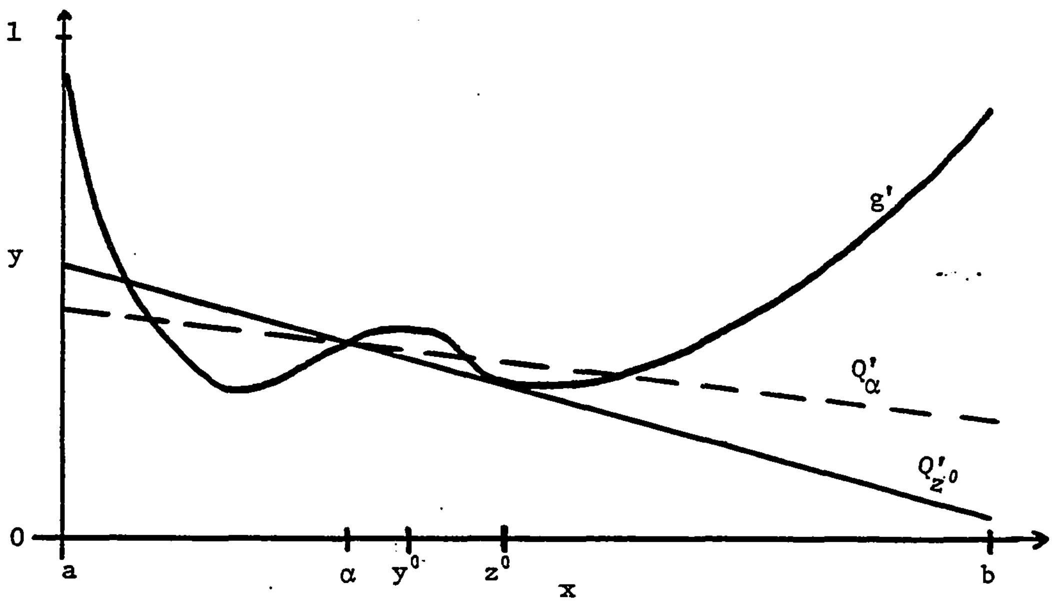


Figure 2.4.3. A third possibility for g'

Hence, whether $g'(y^0) < Q'_{z^0}(y^0)$ or $g'(y^0) \geq Q'_{z^0}(y^0)$,

one may construct an α ,$\alpha < z^0$ (see 2.4.37, 2.4.39), such that $g'(\alpha) = Q'_{z^0}(\alpha)$ (see 2.4.34, 2.4.40) and $\int_a^\alpha g' - Q'_{z^0} < 0$ (see 2.4.38, 2.4.43).

Therefore, considering Q_α

$$\int_a^\alpha g' - Q'_{z^0} + \int_a^\alpha Q'_\alpha - g' < 0 \tag{2.4.44}$$

so that

$$\int_a^\alpha Q'_\alpha - Q'_{z^0} < 0 \ , \tag{2.4.45}$$

and, by S.3,

$$Q'_\alpha\,(z^0) > Q'_{z^0}(z^0) = g'(z^0) \tag{2.4.46}$$

which contradicts 2.4.8.

Since assumptions 2.4.8 and 2.4.30 are incompatible, 2.4.8 implies the complement of 2.4.30 which is precisely 2.4.7.

The following theorems are very similar to Theorem 2.4.1. Proofs are not given since they are so much like the proof of Theorem 2.4.1.

Theorem 2.4.2

Let g be a function satisfying Condition 2.3.1. Then

$$g \leq Q_t \text{ on } [a,b] \quad \forall\, t \in (a,b) \tag{2.4.47}$$

if and only if

$$g' \leq Q_t' \text{ on } [t,b) \quad \forall\, t \in (a,b)\ . \tag{2.4.48}$$

Condition 2.4.1

The function g is finite at b and has continuous derivative on (a,b).

The quadratic function W_t is defined exactly as Q_t in definition 2.3.1 except that

$$W_t(b) = g(b) \tag{2.4.49}$$

rather than $W_t(a) = g(a)$ as in 2.3.21.

Theorem 2.4.3

Let g be a function satisfying Condition 2.4.1. Then

$$g \geq W_t \text{ on } [a,b] \quad \forall\, t \in (a,b) \tag{2.4.50}$$

if and only if

$$g' \geq W_t' \text{ on } (a,t] \quad \forall\, t \in (a,b)\ . \tag{2.4.51}$$

Theorem 2.4.4

Let g be a function satisfying Condition 2.4.1. Then

$$g \leq W_t \text{ on } [a,b] \quad \forall\, t \in (a,b) \tag{2.4.52}$$

if and only if

$$g' \leq W_t' \text{ on } (a,t] \; \forall\, t \in (a,b). \tag{2.4.53}$$

It should be noted that, in Theorems 2.4.1 and 2.4.2, b could be taken as $+\infty$, and, in Theorems 2.4.3 and 2.4.4, a could be taken as $-\infty$. The crucial part of this technique is that one has a finite end point at which g is finite. This is because of the manner in which Q_t and W_t are defined.

Another point worth mention is that Q_t and W_t are defined only for $t \in (a,b)$, whereas certain μ_1, μ_2 combinations would require that $t = a$ or $t = b$ for finite intervals $[a,b]$. However, these cases are not really of interest since either case implies that the distribution is completely specified (i.e., it is either degenerate or has mass only at the end points.) In either case, there is only one distribution satisfying the μ_1, μ_2 constraints so that, being the only feasible distribution, it is therefore optimal. Thus, the specification $t \in (a,b)$. (Note: actually, g satisfying Condition 2.3.1 does not even guarantee g finite at b, e.g., $g = -\ln(1-x)$ on $[0,1]$.)

A very natural question is "What type of function g would satisfy 2.4.8?" Actually, in the following section it will be shown that a rather broad class of functions satisfies 2.4.8.

Other conditions are given which are sufficient for 2.4.7. One of these is convexity of g' ; a condition which is usually not difficult to verify. Additionally, a few results dealing with the nature of Q_t' are also presented.

5. Additional Results

This section presents conditions which are often easier to apply than 2.4.8, i.e., $g' \geq Q_t'$ on $[t,b)$ $\forall\, t \in (a,b)$. Additionally, some characteristics of Q_z are explicitly stated and proved. The purpose is to more fully describe conditions when the mathematical programming approach developed in Sections 2 and 3 of this chapter may be applied.

This is not to say that 2.4.8 is not useful. Rather, more powerful statements may often ease the analyst's task. For example, a result of this section is that convexity of g' implies 2.4.8. So that if one has, say, $g(x) = e^{cx}$, then upper and lower attainable bounds for any finite interval [a,b] quickly follow from 2.3.26. One attainable bound results if only one endpoint is finite. The hope is that, as the cost of complexity decreases, one becomes less willing to pay the price of uncertainty.

Some results concerning the relationship of g' and Q_z' will be presented first. The behavior of Q_t' as a function of t will also be examined, one benefit being aid in later proofs.

The following lemma, besides proving useful for the next theorem, provides some insight concerning the shape of g' for a function g satisfying 2.4.7.

Lemma 2.5.1

Let g satisfy Condition 2.3.1 and 2.4.7. Define

$$s_z = \inf_x \{x|g'(x) > Q_z'(x),\ x \in (z,b]\} \tag{2.5.1}$$

with $s_z = b$ if the set is empty. Then $g' > Q_z'$ on $(s_z,b]$.

Proof: The approach will be to show that, if there exist $z,\ t \in (s_z,b]$ with $g'(t) = Q_z'(t)$, then $g(z) < Q_t(z)$, violating 2.4.7.

Assume that, for some

$$z \in (a,b),\ \exists\ t \in (s_z,b] \ni g'(t) = Q_z'(t) . \tag{2.5.2}$$

(Note: Theorem 2.4.1 guarantees $g' \geq Q_z'$ on $[z,b)$; so only the case of equality need be considered.)

Now, $t > s_z$ and hence, by the definition of s_z in 2.5.1, there is some point r in (s_z,t) such that $g'(r) > Q_z'(r)$. Continuity considerations imply an open interval containing r on which $g' > Q_z'$, i.e., S.4 again.
Therefore,

$$\int_{s_z}^{t} g' - Q_z' > 0 . \tag{2.5.3}$$

Additionally, 2.5.1 means that $g' = Q_z'$ on $[z,s_z]$ so that

$$\int_z^{s_z} g' - Q_z' = 0$$

and therefore,

$$\int_a^{s_z} g' - Q'_z = 0 \text{ .} \tag{2.5.4}$$

From 2.5.3 and 2.5.4 it follows that $\int_a^t g' - Q'_z > 0$.

Of course, $\int_a^t g' - Q'_t = 0$ so that

$$\int_a^t Q'_t - Q'_z > 0 \text{ .} \tag{2.5.5}$$

Appealing once more to S.3, Q'_t and Q'_z are straight lines with intersection at t, so that $Q'_t > Q'_z$ on $[a,t)$.

It immediately follows that

$$\int_a^z Q'_t - Q'_z > 0 \tag{2.5.6}$$

so that

$$\int_a^z Q'_t - g' > 0 \tag{2.5.7}$$

or $Q_t(z) > g(z)$, contradicting the assumption 2.4.7. Hence, the result.

The simple idea of this lemma is that, if g' ever exceeds Q'_z to the right of z, it remains greater. This notion is, of course, present in the proof of Theorem 2.4.1; but, as it concerns the behavior of Q'_z to the right of z relative to g', it seems worthwhile to explicitly state it here.

The next theorem gives a further characterization of Q_t. More than merely a mathematical exercise, a better understanding of Q_t provides insight into the type of function g which will satisfy 2.4.7. Depending on the function g, it may be very difficult to show infeasibility using some of the criteria, but much easier using others. This would appear to be desirable since a useful method should provide rules for disallowing some functions, rather than being confined to the verification of functions which do yield feasible Q_z.

In the theorem, it is useful to recall from 2.3.20 that

$Q_t(x) = \lambda_0 + \lambda_1 x + \lambda_2 x^2$, subject to 2.3.21 - 2.3.23. The coefficients are, of course, functions of a, t, and g. This should be kept in mind in spite of the terse notation which gives no hint of this.

<u>Theorem 2.5.1</u> Let g satisfy Condition 2.3.1. Suppose that

$$g \geq Q_t \text{ on } [a,b] \ \forall\ t \in (a,b). \tag{2.5.8}$$

Then

1) λ_2 is nondecreasing in t, (2.5.9)

2) Q_z' and Q_t', when nonidentical, cross in the interval $(a,\max(z,t))$, and (2.5.10)

3) $Q_t'(a)$ is nonincreasing in t. (2.5.11)

Proof: Choose $t, z \varepsilon (a,b)$ with $t > z$.

By Theorem 2.4.1, $g' \geq Q_z'$ on $[z,b)$.

From 2.3.22 it follows that

$$Q_t'(t) = g'(t) \geq Q_z'(t). \tag{2.5.12}$$

In view of Lemma 2.5.1, the case of equality is uninteresting since Q_t' and Q_z' are then identical functions so that Q_s' remains unchanged on $[z,t]$.

The other case,

$$Q_t'(t) > Q_z'(t) , \tag{2.5.13}$$

also yields the results quite simply. They follow immediately upon showing that Q_t' and Q_z' must cross in the interval (a,t).

If $Q_t' > Q_z'$ on (a,t), i.e., Q_t' and Q_z' do not cross, then

$$\int_a^z Q_t' - Q_z' > 0, \tag{2.5.14}$$

or, more simply,

$$Q_t(z) > Q_z(z) = g(z),$$

violating 2.4.7. Hence, $\exists\ w \ \varepsilon\ (a,t) \ni Q'_z(w) = Q_t'(w)$,

and 2.5.10 is shown. Upon application of S.3 it is readily apparent that 2.5.9 and 2.5.11 are also proven. That is, $2\lambda_2$ is the slope of Q_t'. Statement S.3 shows that Q_t' has greater slope than Q_z' as well as $Q_z'(a) > Q_t'(a)$, so that the theorem is proved.

Another useful, if somewhat obvious, result is that λ_0, λ_1, and λ_2 are continuous functions of t. This is formalized in the following lemma.

Lemma 2.5.2

Let g satisfy Condition 2.3.1. Then the coefficients of Q_t, i.e., λ_0, λ_1, and λ_2 as defined in 2.3.24, are continuous in t.

Proof: The solution of 2.3.24 is

$$\lambda_2 = \frac{-1}{(t-a)^2}\left\{g(t) - g(a) - (t-a)g'(t)\right\} \tag{2.5.15}$$

$$\lambda_1 = g'(t) - 2t\,\lambda_2 \tag{2.5.16}$$

$$\lambda_0 = g(a) - ag'(t) - a(a-2t)\,\lambda_2 \tag{2.5.17}$$

Note that a and $g(a)$ are constants and also that g and g' are continuous on (a,b). Since $t \in (a,b)$, the denominator of 2.5.15 is nonzero and hence λ_2 is continuous, being the quotient of continuous functions. The continuity of λ_0 and λ_1 follows since each is just composed of sums and products of continuous functions.

It will be seen that the continuity of λ_1 and λ_2 plays a key role in the next theorem. This theorem, while seemingly not very easy to apply, nevertheless yields the important convexity condition, stated as Corollary 2.5.

Theorem 2.5.2

Let g satisfy Condition 2.3.1. Furthermore, suppose that, for every $t \in (a,b)$, g' coincides with Q_t' on at most two distinct intervals. Then g lies either entirely above or below Q_t on $[a,b]$.

Proof: The approach will be to show that $g' \geq Q_t'$ or $g' \leq Q_t'$ on $[t,b)$ $\forall$ $t \in (a,b)$ so that Theorem 2.4.1 may be applied.

Consider some $t \in (a,b)$. By definition $g'(t) = Q_t'(t)$, so that t lies in one of the intervals. Call this interval $[\ell_t, r_t]$. If there is another interval of coincidence, call it $[\ell_t', r_t']$.

Now, the constraint

$$\int_a^t g' - Q_t' = 0 \tag{2.5.18}$$

implies that either 1) $a = \ell_t < r_t$ or 2) $[\ell_t', r_t']$ is a subinterval of (a, ℓ_t). The first case is rather obvious and the second is the alternate when g' and Q_t' do not coincide on $[a,t]$ since $g' - Q_t'$ changes sign at most twice by assumption.

Case one will be considered first. If there is no other interval of coincidence, then g' lies entirely to one side of Q_t'. Hence, the task is to show that, if there is another interval, the assumptions are violated.

Suppose that there is an interval of coincidence $[\ell_t', r_t']$ to the right of r_t. Without loss of generality assume that $Q_t' > g'$ on (r_t, ℓ_t'). (2.5.19)

Let L be the line through the point $(a,Q_t'(a))$ lying below g' on (r_t,ℓ_t'), with equality for some point in that interval. Define

$$x^* = \inf_x \{x \mid L(x) = g'(x),\ x \in (r_t,\ell_t')\}. \tag{2.5.20}$$

Figure 2.5.1, illustrates what one possible such situation might look like. Since $L \le g' \le Q_t'$ on (a,ℓ_t') it is clear that

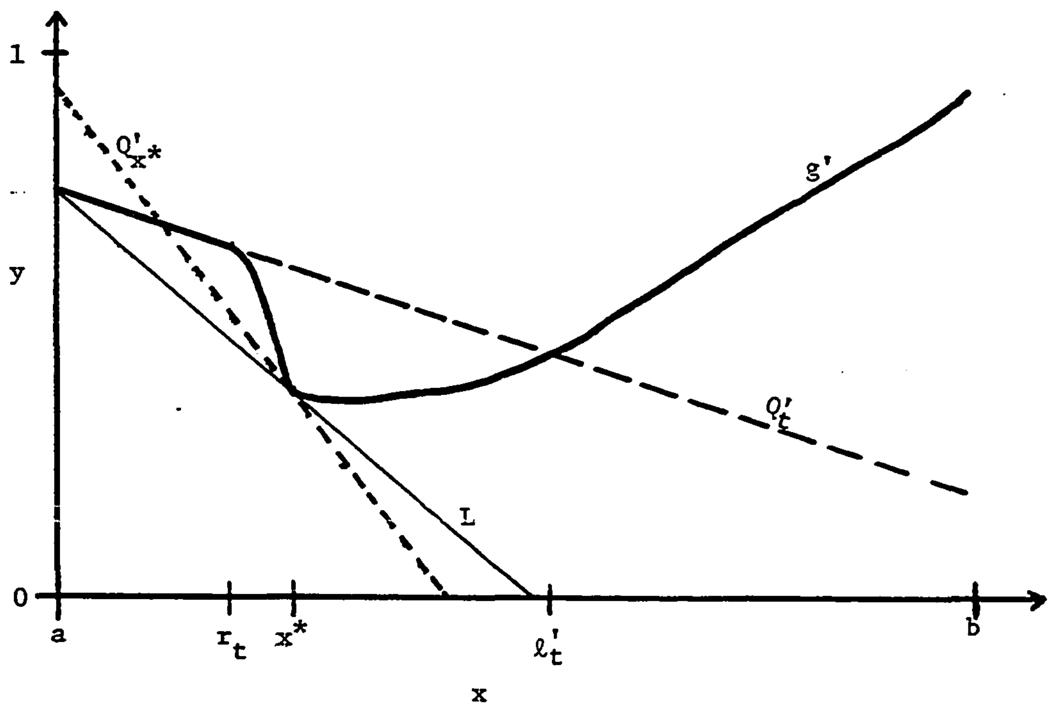


Figure 2.5.1. First illustration for Theorem 2.5.2

$$\text{slope of } Q'_{x^*} < \text{slope of } L < \text{slope of } Q_t'. \tag{2.5.21}$$

Consider any point $z \varepsilon (r_t,x^*)$. Since $Q_t' > g'$ on (r_t,x^*), from 2.5.19, 2.5.20 and

$$\int_a^z g' - Q_z' = 0 \qquad (2.5.22)$$

it follows that Q_t' and Q_z' cross in the interval (a,z). This, of course means that $Q_z'(a) > Q_t'(a)$ and hence $Q_z' > g'$ on some interval (a,w), $w < z$. Of course this means that $g' > Q_z'$ on some interval between w and z and hence distinct intervals of coincidence occur to the left of z and at z. By choosing z near enough to r_t, the continuity of λ_2 and λ_1 ensures a third interval of coincidence between x^* and ℓ_t'. That is, Q_{r_t}' passes through $(\ell_t',g'(\ell_t'))$, Q_{x*}' passes through $(x*,g'(x*))$, g' connects those two points and the slope of Q_{r_t}' exceeds the slope of Q_{x*}'. Since λ_2 is continuous, there are points in $(r_t,x*)$ yielding every possible slope between that of Q_{r_t}' and Q_{x*}'. Therefore, there is some $z* \varepsilon (r_t,x*)$ which crosses g' in the interval $(x*,\ell_t')$. This third crossing violates the assumptions, of course, so that case 1 yields g' lying entirely to one side of Q_t' on $[t,b)$.

Case 2 entails no essential difference. The continuity of λ_1 and λ_2 again plays a key role. This time without loss of generality let $g' > Q_t'$ on (a,ℓ_t') and $g' < Q_t'$ on (r_t',ℓ_t) i.e. the constraint $\int_a^t g' - Q_t' = 0$ assures that $g' - Q_t'$ has opposite signs on the two intervals of noncoincidence to the left of t.

Assume that $g' < Q_t'$ on (r_t, b). Note that otherwise $g' \geq Q_t'$ on (r_t, b), which is the result to be shown.

Now let L be the line passing through $(r_t, g'(r_t))$ such that $L \leq g'$ on (r_t', ℓ_t) with at least one point of equality. (2.5.23)

One possible such configuration is pictured in Figure 2.5.2.

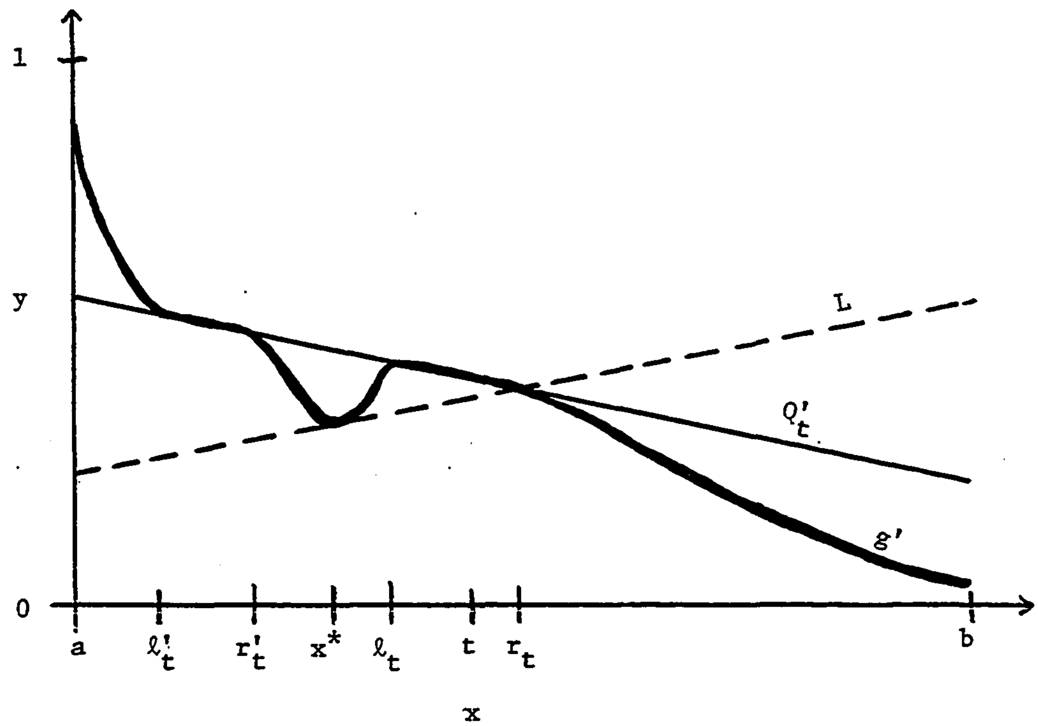


Figure 2.5.2. Second illustration for Theorem 2.5.2

Define

$$x^* = \inf_x \{x \mid L(x) = g'(x),\ x \in (r_t', \ell_t)\} . \tag{2.5.24}$$

Now, since λ_1 and λ_2 are continuous functions of the parameter z for Q_z, then $Q_z'(x) = 2\lambda_2 x + \lambda_1$ is also a continuous function of z. Hence, by choosing ε small enough, one can find $z^* \varepsilon (r_t, r_t + \varepsilon)$ such that Q_{z*}' crosses g' on (a,x^*), on (x^*,r^t), and at $z^* > r^t$, i.e., three distinct intervals of coincidence, since the intervals of coincidence are subintervals of nonoverlapping intervals.

The key to doing this, of course, is choosing ε small enough that the quantity

$$\int_{r_t}^{r_t+\varepsilon} Q_t' - g'$$

is very small.

This forces Q_{z*}' and Q_t' to cross in (a,r_t), so that Q_{z*}' crosses g' in (x^*,r_t) since Q_{z*}' has to lie above $L(x^*)$. Of course ε may also be made small enough that $Q_{z*}'(a)$ is close to $Q_t'(a)$, forcing a crossing in (a,x^*).

Since the assumption of two intervals of coincidence is violated, the assumption $g' < Q_t'$ on (r_t,b) is shown to lead to a contradiction. Hence, case 2 also implies that g' lies entirely to one side of Q_t' on $[t,b)$, and, in this case, on the "proper" side, i.e., the sign of $g' - Q_t'$ is the same on (a,ℓ_t') and on (r_t,b).

Only "oscillating" behavior could be of concern now. That is, $g' \geq Q_t'$ on $[t,b)$ for some t's, and $g' \leq Q_t'$ on $[t,b)$ for other t's. Continuity considerations for λ_1 and λ_2, as well as the method of proof, show this not to be the case, however, since this would lead to multiple crossings. This completes the argument.

The proof stands completed then. It has been shown that, if g satisfies condition 2.3.1, and, for every $t \in (a,b)$, g' coincides with Q_t' on at most two distinct intervals, then, for every $t \in (a,b)$, g lies either entirely above or entirely below Q_t on $[a,b]$.

Two important corollaries immediately follow since a straight line intersects a concave or convex function at most twice.

Corollary 2.5.2 Let g satisfy Condition 2.3.1. Then, if g' is convex, $g \geq Q_t$ on $[a,b]$ for every $t \in (a,b)$.

Corollary 2.5.3 Let g satisfy Condition 2.3.1. Then, if g' is concave, $g \leq Q_t$ on $[a,b]$ for every $t \in (a,b)$.

A very simple proof of these corollaries exists independently of Theorem 2.5.2. However, they seem to follow rather naturally here.

The condition g' convex or concave is fairly strong, but not at all uncommon. The following theorem generalizes the idea, so that a g' function which is "piecewise convex" subject to certain other conditions may also yield feasibility.

Theorem 2.5.3

Let g satisfy Condition 2.3.1. Furthermore, suppose that, for some $c \in (a,b)$, g satisfies the following conditions:

1) $g' \geq Q_t'$ on $[t,c]$, $\forall\, t \in (a,c)$ (2.5.25)

2) $g' \geq Q_c'$ on $[c,b)$ (2.5.26)

3) g' is convex on $[c,b)$ (2.5.27)

Then $g \geq Q_t$ on $[a,b]$ $\forall\, t \in (a,b)$.

Proof: The first condition 2.5.25 implies that $g \geq Q_t$ on $[a,c]$, $\forall\, t \in (a,c]$, by Theorem 2.4.1 (Note that $t = c$ will be permitted since $c < b$.) Applying Theorem 2.5.1 and S.3, Q_c' lies above Q_t' on $[c,b)$ since Q_c' has greater slope and their point of intersection is in the interval $(a,c]$. Hence,

$g' \geq Q_c' \geq Q_t'$ on $[c,b)$ for $t \in (a,c]$ and therefore

$g' \geq Q_t'$ on $[t,b)$ for $t \in (a,c]$.

It now only remains to show that $g' \geq Q_t'$ on $[t,b)$ for $t \in (c,b)$. Choose some $t \in (c,b)$ and let ℓ_t be a subgradient of g' at t. Furthermore, let L be a line through the point $(t,g'(t))$ with slope ℓ_t. Then, by the convexity of g' on $[c,b]$, $g' \geq L$ on $[c,b]$. Of course, L and Q_t' intersect at the point $(t,g'(t))$. If L passes through the point $(c,g'(c))$ then convexity implies that g' is linear on $[c,t]$, so that $Q_t' = Q_c'$ and hence $g' \geq Q_t'$ on $[t,b)$. Otherwise, $L(c) < g'(c)$ and $L(t) > Q_c'(t)$ so that L and Q_c' cross in (c,t). By S.3, $Q_c' > L$ on $[a,c]$ so that

$$\int_a^c Q_c' - L > 0, \tag{2.5.28}$$

which implies

$$\int_a^c g' - L > 0. \tag{2.5.29}$$

Recalling that convexity of g' on $[c,b]$ and $L(c) < g'(c)$ imply that

$$\int_c^t g' - L > 0 , \tag{2.5.30}$$

it follows that

$$\int_a^t g' - L > 0 . \tag{2.5.31}$$

The result. is that

$$\int_a^t Q_t' - L > 0 , \tag{2.5.32}$$

so that application of S.3 yields

$$g' \geq L > Q_t' \quad \text{on} \quad (t,b) .$$

Hence, $g' \geq Q_t'$ on $[t,b)$ $\forall$ $t \in (a,b)$ so that Theorem 2.4.1 applies and $g \geq Q_t$ on $[a,b]$ $\forall$ $t \in (a,b)$.

The theorem at once yields the following corollary.

<u>Corollary 2.5.4</u> Let g satisfy Condition 2.3.1. Suppose that there exists $\{a_i | a = a_0 < a_1 < \ldots < a_k = b, i \in I = \{1, 2, \ldots, k\}\}$ such that

1) $g' \geq Q_t'$ on $[t,a_1]$, $\forall$ $t \in (a,a_1)$, (2.5.33)

2) g' is convex on $[a_i,a_{i+1})$, $\forall$ $i \in I$, (2.5.34)

3) $g' \geq Q_{a_i}'$ on $[a_i,a_{i+1})$, $\forall$ $i \in I$ (2.5.35)

Then $g \geq Q_t$ on $[a,b]$ $\forall$ $t \in (a,b)$.

which implies

$$\int_a^c g' - L > 0. \tag{2.5.29}$$

Recalling that convexity of g' on $[c,b]$ and $L(c) < g'(c)$ imply that

$$\int_c^t g' - L > 0 , \tag{2.5.30}$$

it follows that

$$\int_a^t g' - L > 0 . \tag{2.5.31}$$

The result is that

$$\int_a^t Q_t' - L > 0 , \tag{2.5.32}$$

so that application of S.3 yields

$$g' \geq L > Q_t' \quad \text{on} \quad (t,b) .$$

Hence, $g' \geq Q_t'$ on $[t,b)$ $\forall$ $t \in (a,b)$ so that Theorem 2.4.1 applies and $g \geq Q_t$ on $[a,b]$ $\forall$ $t \in (a,b)$.

The theorem at once yields the following corollary.

<u>Corollary 2.5.4</u> Let g satisfy Condition 2.3.1. Suppose that there exists $\{a_i | a = a_0 < a_1 < \ldots < a_k = b, i \in I = \{1, 2, \ldots, k\}\}$ such that

1) $g' \geq Q_t'$ on $[t,a_1]$, $\forall$ $t \in (a,a_1)$, (2.5.33)

2) g' is convex on $[a_i,a_{i+1})$, $\forall$ $i \in I$, (2.5.34)

3) $g' \geq Q_{a_i}'$ on $[a_i,a_{i+1})$, $\forall$ $i \in I$ (2.5.35)

Then $g \geq Q_t$ on $[a,b]$ $\forall$ $t \in (a,b)$.

The proof of this is simply a recursive application of Theorem 2.5.3 and will not be given here.

Though there are undoubtedly other sufficient conditions for $g \geq Q_t$, those given here form a fairly broad basis. Section 4 of Chapter 3 presents the assumptions of Mantell which are sufficient to insure a lower bound, and Chapter four briefly presents others. At this time, a more profitable course is to proceed to the last section of this chapter where elements of preceeding sections are drawn together to form some simply stated results which address the objectives of this thesis.

6. Summation

Before concluding this chapter, it is advantageous to present a theorem which deals directly with the solution of P_1. Additionally, some details are presented which state succinctly the ideas presented previously.

The first feature of interest is that the methodology presented in this work requires a finite endpoint. This presentation does not address the problem P_1 when (a,b) is $(-\infty, +\infty)$. In fact, as indicated earlier, it is not helpful to consider an endpoint as finite but "heading toward infinity" since the bound 2.3.26 becomes the Jensen bound in the limit if $g(x) = o(x^2)$. Hence, to effectively use the methodology of this thesis one must have at least one finite endpoint for the interval of integration. This is apparent upon examination of the manner in which Q_z is defined.

The following theorem, while not providing a complete solution to P_1, nevertheless does solve P_1 for a "reasonably large" class of functions g.

Theorem 2.6.1

Let g be a real valued function on the interval $[a,b]$, a finite. Furthermore let g be continuous at a and have continuous derivative, g', on (a,b).

Then

$$\frac{g(a)(\mu_2-\mu_1^2) + g\left(\frac{\mu_2-a\mu_1}{\mu_1-a}\right)(\mu_1-a)^2}{(\mu_1-a)^2 + (\mu_2-\mu_1^2)} \equiv L(\mu_1,\mu_2,g,a) \tag{2.6.1}$$

is the unique lower attainable bound for

$$\int_a^b g\,dF , \tag{2.6.2}$$

where

$$F \in \mathcal{F} = \{F \mid \int_a^b x\,dF = \mu_1,\ \int_a^b x^2 dF = \mu_2,\ F \text{ is a cdf on } [a,b]\}, \tag{2.6.3}$$

if any of the following conditions is met:

1. $g'(x) \geq g'(t) - \frac{2(x-t)}{(t-a)^2}(g(t)-g(a)-(t-a)g'(t))$

$$\forall\ t \in (a,b)\ ,\ x \in [t,b) \tag{2.6.4}$$

2. g' is convex on $[a,b]$ (2.6.5)

3. $\exists \{a_i \mid a = a_0 < a_1 < \ldots < a_k = b\} \ni g' \geq Q_t'$ on $[t,a_1)$ $\forall\ t \in (a,a_1)$, g' is convex on $[a_i,a_{i+1})$,

$i = 0, 1, \ldots, k\text{-}1$, and $g' \geq Q'_{a_i}$ on $[a_i, a_{i+1})$,
$i = 0, 1, \ldots, k\text{-}1$. (2.6.6)

4. For all $t \in (a,b)$, g' coincides with Q'_t on at most two distinct (possibly degenerate) intervals and, for some t , $g'(a) > Q'_t(a)$, (2.6.7)

where Q_t is as in Definition 2.3.1 together with $Q_t(a) = g(a)$.

Proof:

Theorems 2.4.1 and 2.5.2 and Corollaries 2.5.2 and 2.5.4 show that any of conditions 1 through 4 implies $g \geq Q_t$ on $[a,b]$. Note that the RHS of 2.6.4 is $Q'_t(x)$. Furthermore, $g(a) = Q_t(a)$ and $g(t) = Q_t(t)$.

Hence, let

$$F^*(x) = \begin{cases} 0 & x < a \\ \dfrac{\mu_2-\mu_1^2}{(\mu_1-a)^2} & a \leq x < \dfrac{\mu_2-a\mu_1}{\mu_1-a} \\ 1 & x \geq \dfrac{\mu_2-a\mu_1}{\mu_1-a} \end{cases} \tag{2.6.8}$$

and $Q^* = Q_{\frac{\mu_2-a\mu_1}{\mu_1-a}}$.

The three essential ingredients are now present for the programming approach. That is, $g \geq Q^*$ on $[a,b]$, $F^* \in \mathcal{F}$, and $g = Q^*$ at the points of increase of F^* .

So, for all $F \in \mathcal{F}$,

$$\begin{aligned} \int_a^b g\,dF &\geq \int_a^b Q^* dF \\ &= \int_a^b Q^* dF^* \qquad (2.6.9) \\ &= \int_a^b g\,dF^* = L(\mu_1,\mu_2,g,a) \end{aligned}$$

The importance of this theorem is that the "shape" of g determines whether the bound 2.6.1 is valid. It depends in no way upon μ_1 or μ_2. Thus, when problem P_1 is posed, one must check g for the right shape, assuming that the problem is feasible, of course. If g does satisfy any of 2.6.4-2.6.7, then the bound 2.6.1 applies. If not, an argument specific to the particular (μ_1,μ_2,g) triple will need to be brought to bear. Chapter four discusses this problem at greater length, elaborating on some of the reasons that a solution such as 2.6.1 does not appear to be possible in the case of general g.

Another point is that this methodology may also be used for upper bounds. If $-g$ is considered rather than g, then the inequalities in Theorem 2.6.1 reverse and "convex" becomes "concave". This was indicated by Theorem 2.4.2. The complete development is parallel and will not be given here.

If b is finite, it may be the case that b may be used as a bound. The development is, again, parallel as indicated by Theorems 2.4.3 and 2.4.4. The following theorem parallels the preceding one.

Theorem 2.6.2

Let g be a real valued function on the interval $[a,b]$, b finite. Let g be continuous at b with continuous derivative, g', on (a,b). Then

$$\frac{g(b)(\mu_2-\mu_1^2) + g\left(\frac{\mu_2-b\mu_1}{\mu_1-b}\right)(\mu_1-b)^2}{(\mu_1-b)^2 + (\mu_2-\mu_1^2)} \equiv L(\mu_1,\mu_2,g,b) \tag{2.6.10}$$

is the unique upper attainable bound for

$$\int_a^b g\,dF ,$$

where

$$F \in \mathcal{F} = \{F \mid \int_a^b x\,dF = \mu_1,\ \int_a^b x^2 dF = \mu_2,\ F \text{ is a cdf on } [a,b]\}, \tag{2.6.11}$$

if any of the following conditions is met:

1. $g'(x) \geq g'(t) - \frac{2(x-t)}{(t-b)^2}(g(t)-g(b)-(t-b)g'(t))$

   $$\forall\ t \in (a,b)\ ,\ x \in (a,t] \tag{2.6.12}$$

2. g' is convex on $[a,b]$ (2.6.13)

3. $\exists \{a_i \mid a = a_0 < a_1 < a_2 < \ldots < a_k = b\}$

   $\ni g' \geq W_t'$ on $(a_k,t]$ $\forall\ t \in (a_k,b)$, g' is convex

   on $[a_i,a_{i+1})$, $i = 0, 1, \ldots, k-1$, and $g' \geq W_{a_{i+1}}'$

   on $(a_i,a_{i+1}]$, $i = 0, 1, \ldots, k-1$. (2.6.14)

4. For all $t \in (a,b)$, g' coincides with W_t' on at most two distinct (possibly degenerate) intervals, and, for some t ,

$$g'(b) > W_t'(b) \; , \tag{2.6.15}$$

where W_t is as defined at 2.4.49.

The proof of this theorem will not be given, but some of the differences with Theorem 2.6.1 will be briefly noted. The chief difference is the use of b throughout where a was used formerly. The other salient difference is that the critical interval endpoints are now the right endpoints, rather than the left. As with Theorem 2.6.1, reversing the inequalities and changing "convex" to "concave" yields a lower, rather than an upper, bound.

It should be noted that only the second condition is common to both theorems. In both cases the first condition is the most general, being implied by those conditions following it. But it is entirely possible for a function, g, on a finite interval, [a,b], to satisfy one of 2.6.4 or 2.6.12 without satisfying the other. Hence, this method does not necessarily generate both an upper and a lower bound whenever it generates either of the two. The second condition, however, gives both, so that the following results.

<u>Theorem 2.6.3</u>

Let g be a real valued continuous function on the finite interval [a,b] with convex derivative, g' . Then for

$$F \in \mathcal{F} = \{F | \int_a^b x dF = \mu_1,\ \int_a^b x^2 dF = \mu_2,\ F \text{ is a cdF on } [a,b]\}$$

one has the inequality

$$L(\mu_1,\mu_2,g,a) \leq \int_a^b g dF \leq L(\mu_1,\mu_2,g,b) \qquad (2.6.16)$$

This result follows immediately from the preceding two theorems of this section. Note that for concave g' the inequalities in 2.6.16 are reversed.

Chapter four will contain several examples to clarify ideas presented here. Preceding that, chapter three will present an application of this method to utility theory.

III. APPLICATIONS TO UTILITY THEORY

1. Introduction

The ideas developed in the previous chapter, though originally motivated by utility theory considerations, are not, however, restricted to utility theory, being more widely applicable. Nevertheless, these ideas easily yield fruitful results for utility analysis. It is the aim of this chapter to relate the implications of chapter two to current utility theory.

Section two presents some general results and bounds for commonly used utility functions. The next section examines approximation to expected utility.

The work of Mantell, and that of Levy and Markowitz, is examined in the light of Chapter two, respectively in the fourth and fifth sections.

2. General Applications

Some rather popular utility functions are presented in this section and are shown to be "of the right shape" for the applicability of Theorem 2.6.1. It will be shown that this will always be the case when certain "reasonable" assumptions are met.

Fishburn and Vickson (1978) summarize several results of E-V analysis and of stochastic dominance. Listed below is their compilation of some common utility functions u which possess certain desirable characteristics. Those properties are that u, u', and u'' are continuous on the domain I of u, with $u' > 0$ and $u'' < 0$. Such a utility function indicates increasing strength of preference and an aversion to risk. This is the type of behavior commonly encountered in many applications.

The functions given by Fishburn and Vickson (1978) are:

$$u_1(x) = -1/(x+d)^c,\ c > 0,\ d > 0,\ I = [a,b] \subset (-d,+\infty) \tag{3.2.1}$$

$$u_2(x) = (x+d)^{1-c},\ c \in (0,1),\ I = [a,b] \subset (-d,+\infty) \tag{3.2.2}$$

$$u_3(x) = -e^{-cx},\ c > 0,\ I = [a,b] \tag{3.2.3}$$

$$u_4(x) = x - cx^2,\ c > 0,\ I = [a,1/(2c)],\ a \in (-\infty,1/(2c)) \tag{3.2.4}$$

$$u_5(x) = \log_e(x+d),\ d > 0,\ I = [a,b] \subset (-d,+\infty) \tag{3.2.5}$$

It is readily apparent that condition 2.3.1 is satisfied for these functions. Straightforward calculation easily yields the third derivatives as:

$$u_1'''(x) = c(c+1)(c+2)/(x+d)^{c+3} \tag{3.2.6}$$

$$u_2'''(x) = c(1-c)(1+c)/(x+d)^{c+2} \tag{3.2.7}$$

$$u_3'''(x) = c^3 e^{-cx} \tag{3.2.8}$$

$$u_4'''(x) = 0 \tag{3.2.9}$$

$$u_5'''(x) = 2/(x+d)^3 \tag{3.2.10}$$

The reader is reminded that the aim throughout is to bound expected utility, as a function of the first two movements, μ_1 and μ_2, of the random payoff. Clearly u_4 is not of interest in this regard, since its expectation is determined by μ_1 and μ_2 .

The remaining four functions all have positive third derivative. This may be readily apparent to utility theorists, of course, since each function exhibits decreasing absolute risk aversion, which always implies $u''' > 0$. Hence, each of these functions has a convex derivative and thereby satisfies Theorem 2.6.1. In fact, Theorem 2.6.3 is applicable so that upper and lower bounds for $\int_a^b u_i dF$ for $F \in \mathcal{F}$, are shown in 2.6.16.

Since utility functions are unique only up to changes in location and scale, it is common to define increasing utility functions to lie on the unit interval with $u(0) = 0$ and $u(1) = 1$. Denoting the transformed functions by $\bar{u}_1$, $\bar{u}_2$, ..., $\bar{u}_5$, one may write the bounds as follows. The lower bound for the expectation over the ith utility function is given by

$$L(\mu_1,\mu_2,\bar{u}_i) = \frac{\mu_1^2}{\mu_2} \bar{u}_i (\mu_2/\mu_1) \quad . \tag{3.2.11}$$

The corresponding upper bound is likewise written as

$$U(\mu_1,\mu_2,\bar{u}_i) = \frac{(\mu_2-\mu_1^2) + \bar{u}_i(\frac{\mu_2-\mu_1}{\mu_1-1})(\mu_1-1)^2}{(\mu_2-\mu_1^2) + (\mu_1-1)^2} \quad . \tag{3.2.12}$$

One of the benefits of scaling to [0,1] is being able to make direct comparisons for different utility functions and distribution function classes. The difference of 3.2.12 and 3.2.11 above provides a measure of uncertainty in approximating $\int_0^1 \bar{u}_i dF$ for $F \in \mathcal{F}$.

That is, the range of possible approximations is reduced from the interval $[\mu_1, \bar{u}_i(\mu_1)]$ (using respectively a lower linear and upper Jensen bound) to the interval $[L(\mu_1,\mu_2,\bar{u}_i), U(\mu_1,\mu_2,\bar{u}_i)]$. Numerical work has shown this to be a substantial improvement, the upper and lower bound often agreeing to at least two or three digits.

For a given utility function, u, the possible range (L,U) for $\int udF$ is computed in terms of μ_1 and μ_2. For example, suppose one's utility function is $u(x) = \log_e(x+1)/\log_e(2)$ on [0,1]. If it is known that a payoff has mean .5 and variance .05, it is easily ascertained that $\int_0^1 udF$ must lie in the interval [.565, .571], giving an order of magnitude reduction in uncertainty regarding $\int udF$, over the first-moment bounding pair [.500, .585]. Furthermore, one may "out-of-hand" dismiss any sure-thing payoff with an utility of less than .565 since F must surely be preferrable. Thus, these bounds could be useful for screening candidates, eliminating those that must be dominated.

Extending this idea further, suppose in the previous example that a payoff distribution, F^*, with mean .51 and variance .055 is available. Traditional E-V analysis ([Markowitz 1952], [Baron 1977]) does not lend itself to comparing these distributions since F^* has larger mean and variance than does F. However, the expected utility $\int_1^0 udF^*$ must be at least .573 since that is the lower bound for distributions with mean .51 and variance .055. The distribution F^* is preferred to F, therefore, since its worst possible expected utility exceeds F's best possible expected utility.

Where it is possible to apply these bounds, this method compares favorably with derivative-based approximators, which in fact may yield values outside the interval of possibilities. As indicated earlier, decreasing absolute risk aversion always implies $u''' > 0$, i.e., convex derivative, u'. Hence, in view of Theorem 2.6.3, the bounds apply whenever one has a finite interval of payoff possibilities and decreasing absolute risk aversion. Since this is the usual case, these bounds may usually be applied.

Some further questions are covered in the remainder of this chapter, some of these in chapter four, and others will be deferred to future research. The first of those questions, already hinted at in the example, is covered in the next section.

3. Approximation

This section presents an approximation for $E_F[u] = \int_a^b u dF$ which, though rather trivial, possesses the property of never being "too far wrong". This present work has been concerned with determining what uncertainty there is when approximating $E_F[u]$ with only the first two moments of F known. The problem of approximation has not yet been directly addressed here, but it seems appropriate to do so now.

It should be recalled that the purpose of utility theory is to simplify the decision maker's task by quantifying each alternative action so that those actions may be ranked in order of preference. The quantity $E_F[u]$ is the numerical index assigned to F. Since F

is often known only to its first two moments, this value is usually approximated by some function of those two moments. This section and the last take up this issue, though we note that this section is not proposing approximation for utility analysis, but rather is suggesting guidelines for approximation if approximation is to be done.

The approximation problem may then be viewed in the following terms. An investor has a utility function, u, on the range [a,b] and is considering an alternative which has a payoff distribution, F, on [a,b] for which only the first two moments are known. How is the value $E_F[u]$ best approximated?

There is no problem if u is linear or quadratic, but these are usually inappropriate forms for u. Fortunately, u' is often convex so that an upper bound, U, and a lower bound, L, are easily obtained in many situations. Barring further information, F might be the distribution yielding the lower bound, the distribution yielding the upper bound, or a distribution yielding a value somewhere between those bounds. One can simply say no more with this limited information.

In view of this, it seems that any reasonable approximation of $E_F[u]$ should lie in the interval [L, U]. There would appear to be no compelling reason to prefer any particular point of this interval for the approximation. However, there are some rather obvious possibilities which suggest themselves.

The first of these is L . This is a rather conservative course since only the worst possible distribution will yield this value. This may, perhaps, be too conservative since a risk averse attitude would already be "built into" the utility curve. If one desires an approximation for use for a maxi-min rule, however, then this approximation would be appropriate (cf Mantell (1976)).

Conversely, one might wish to use U. A rather optimistic approach would be to select the alternative having the greatest potential for gain. This is probably not a good procedure, but it should be pointed out that some derivative-based approximations may often fall near or even exceed U . Hence, U often differs little from current practice, and in fact provides an additional degree of protection.

A more reasonable course would be to take $m = (U + L)/2$ as the required approximation. This has the desirable property of never being more than $(U - L)/2$ away from $E_F[u]$ i.e.,

$$|m - E_F[u]| \leq (U - L)/2 \ . \tag{3.3.1}$$

Clearly, no other approximator possesses this mini-max property. In fact, an approximator could do a great deal worse, especially if it falls outside the interval [L, U]. Furthermore, it has the advantage of utilizing information from both L and U . This could be quite important in some applications.

For example, suppose an investor has a choice between two alternatives A and B and finds bounds for $E_{F_A}[u]$ and $E_{F_B}[u]$. If the bounds for the former are [.5, .65] and for the latter are [.51, .52], then the first would seem to present a superior choice unless there were a dire need for a guarantee of at least .51 .

The midpoint approximator would appear to be quite suitable for general use, but, as with anything, should not be used blindly. Its use, however, should be preferred to any approximator lacking boundary guarantees. The same could actually be said for any point of the interval $[L, U]$, and there appears at this time to be no definitive choice. The last section will examine how this work can be used to evaluate the approximator of Levy and Markowitz.

4. Relation to the Work of Mantell

In this section, the results of Mantell (1976) are presented and shown to follow as a consequence of Theorem 2.6.1. Mantell's paper is one of the few efforts in the literature to find bounds for expected utility when the distribution under consideration is only partially known. In view of this, it is surprising that this paper has been largely ignored. Indeed, this present work might not have been undertaken were not the Mantell paper so little known, going virtually unnoticed in the literature.

That paper considers the problem of selecting a portfolio. An investor has a choice among portfolios, each portfolio being described by a vector from the set $W = \{w | w = (w_1, w_2, \ldots, w_n), w_i \geq 0 \; \forall_i, \Sigma w_i = 1\}$. The return $r = \sum_i w_i r_i$ for the portfolio is of interest where r_i is the random return of the ith asset.

Mantell proposed the use of the maxi-min rule of choosing the available portfolio having the largest lower bound. The rule was stressed rather than the bound or the method of obtaining the bound.

This rule, which does provide a degree of protection, is not completely satisfactory in that a potentially large gain may be bypassed in favor of an alternative which is only marginally safer. This shortcoming was pointed out at the close of the previous section.

Two propositions were presented in order to apply the maxi-min rule. For both of these it was assumed that the utility function, u, was strictly monotone increasing, strictly concave, and with $u(0) = 0$. The first proposition uses only the first moment. That first result is

<u>Proposition 1</u> If each individual asset's rate of return r_i is assumed to lie in a finite interval $[0, r_i^*]$, then

$$E[u(r)] \geq \frac{E(r)\ u(r^*)}{r^*} \tag{3.4.1}$$

where $r^* = \sum_i w_i r_i^*$.

A simple application of the programming approach of Chapter II yields this result. Instead of a quadratic, one need only consider the line L passing through $(0, 0)$ and $(r^*, u(r^*))$ together with the lottery which yields r^* with probability $E(r)/r^*$ and 0 with probability $1 - E(r)/r^*$. Since $u \geq L$ on $[0, r^*]$, it follows that

$$E[u(r)] \geq \frac{E(r)}{r^*} u(r^*) + (1 - \frac{E(r)}{r^*}) u(0)$$

$$= \frac{E(r)u(r^*)}{r^*}$$

Mantell's proof hinged on the monotone non-increasing property of the function $u(r)/r$. That approach, however, does not demonstrate the attainability of the bound. The other result, utilizing knowledge of a second moment to obtain an improved bound, is given by

Proposition 2 If the investor's average utility function, $u(r)/r$, is convex for all $r \varepsilon (0, \infty)$, then

$$E[u(r)] \geq \frac{\mu^2}{\mu^2 + \sigma^2} u(\mu + \frac{\sigma^2}{\mu}) \tag{3.4.2}$$

where $\mu = E(r)$ and $\sigma^2 = Var(r)$.

It will now be shown that this proposition follows from Theorem 2.6.1.

Proof: Convexity of $u(r)/r$ implies that

$$\frac{d}{dr}\left(\frac{u(r)}{r}\right) = \frac{ru'(r) - u(r)}{r^2} \tag{3.4.3}$$

is monotone non-decreasing (actually, increasing since u is strictly concave). Consider some $z \varepsilon (0, \infty)$. Then for $x \varepsilon (z, \infty)$ it follows that

$$\frac{d}{dr}\left(\frac{u(r)}{r}\right)\Bigg|_{r=x} \geq \frac{d}{dr}\left(\frac{u(r)}{r}\right)\Bigg|_{r=z} \tag{3.4.4}$$

which is

$$\frac{xu'(x) - u(x)}{x^2} \geq \frac{zu'(z) - u(z)}{z^2} . \tag{3.4.5}$$

Recall from 2.5.15 that for Q_z one has

$$\lambda_2 = \frac{-1}{(z-0)^2}\{u(z) - u(0) - (z-0)\,u'(z)\} \tag{3.4.6}$$

$$= \frac{-1}{z^2}\{u(z) - zu'(z)\} \tag{3.4.7}$$

$$= \frac{zu'(z) - u(z)}{z^2}\,, \tag{3.4.8}$$

i.e., $\frac{d}{dr}\left(\frac{u(r)}{r}\right)\Big|_{r=z}$ is actually the leading coefficient of Q_z .

Therefore, 3.4.5 becomes

$$\frac{xu'(x) - u(x)}{x^2} \geq \lambda_2 \tag{3.4.9}$$

or

$$u'(x) \geq \frac{u(x)}{x} + \lambda_2 x,\ x \in (z, \infty)\ . \tag{3.4.10}$$

Now, since $u(r)/r$ is convex, consider the support of $u(r)/r$ at z . Convexity yields

$$\frac{u(x)}{x} \geq \frac{u(z)}{z} + \lambda_2(x-z)\ \forall\ x \in (0, \infty)\ . \tag{3.4.11}$$

Combining 3.4.10 and 3.4.11 it follows that

$$u'(x) \geq \frac{u(z)}{z} + \lambda_2(x-z) + \lambda_2 x \quad \forall\ x \in (z, \infty) \tag{3.4.12}$$

$$= 2\lambda_2 x + \left(\frac{u(z)}{z} - z\lambda_2\right) \tag{3.4.13}$$

$$= 2\lambda_2 x + \left\{\frac{u(z)}{z} + z\lambda_2 - 2z\lambda_2\right\} \quad . \tag{3.4.14}$$

Now, upon substitution from 3.4.8

$$\text{RHS} = 2\lambda_2 x + \left\{\frac{u(z)}{z} + z\,\frac{zu'(z) - u(z)}{z^2} - 2z\lambda_2\right\} \tag{3.4.15}$$

$$= 2\lambda_2 x + (u'(z) - 2z\lambda_2) \tag{3.4.16}$$

$$= 2\lambda_2 x + \lambda_1 \tag{3.4.17}$$

since, from 2.5.16 $\lambda_1 = u'(z) - 2z\lambda_2$. This is recognizable, of course as $Q'_z(x)$. Hence, the conditions of Proposition 2 imply that

$$u'(x) \geq Q'_z(x) \quad \forall\, x \in (z, \infty) \ . \tag{3.4.18}$$

Theorem 2.4.1 or Theorem 2.6.1 may now be applied so that the bound from 2.3.26 is valid and

$$E[u(r)] \geq \frac{u(0)(\mu_2-\mu_1^2) + u(\frac{\mu_2-0\ \mu_1}{\mu_1-0})(\mu_1-0)^2}{(\mu_1-0)^2 + (\mu_2-\mu_1^2)} \tag{3.4.19}$$

which simplifies to

$$E[u(r)] \geq \frac{\mu_1^2}{\sigma^2 + \mu_1^2}\, u(\frac{\sigma^2 + \mu_1^2}{\mu_1}) \tag{3.4.20}$$

which is the result that was to be proved. Therefore, the results of Mantell are shown to be special cases of the more general method of Chapter two.

In summary, then, this section has presented Mantell's results and shown them to be consequences of results in Chapter II of this present work. Thus, those earlier results have been considerably strengthened and extended.

5. Impact on the Levy-Markowitz Approximator

In this section, current mean-variance approximation is considered. The results of the previous chapter will be used to examine the most widely-used approximator. The purpose will be to simply evaluate whether for many common utility functions, the approximator gives "reasonable" results.

Levy and Markowitz introduced a class of approximators to the readers of The American Economic Review in 1979. The class is characterized by the parameter k in the formulation

$$f_k(\mu,\sigma^2,u) = u(\mu) + \frac{u(\mu+k\sigma) + u(\mu-k\sigma) - 2u(\mu)}{2k^2} \tag{3.5.1}$$

The class contains the popular approximation based on the Taylor series about μ

$$u(\mu) + \frac{\sigma^2}{2} u''(\mu) , \tag{3.5.2}$$

whose relation to the Taylor series is seen upon defining $f_0(\mu,\sigma^2,u)$ as $\lim_{k \to 0} f_k(\mu,\sigma^2,u)$, and appealing to L'Hospital's rule. It also includes the secant approximation

$$\frac{u(1 + \mu + \sigma) + u(1 + \mu - \sigma)}{2} \tag{3.5.3}$$

for $k = 1$. Not included is the approximation based on the Taylor series about zero

$$u(0) + u'(0)\mu + \frac{\mu^2 + \sigma^2}{2} u''(0) , \qquad (3.5.4)$$

which is the only other approximator to receive much attention. This latter is usually a rather poor approximator, however, so that class f_k suffers little for lack of it. Loistl (1976) presents some problems encountered when using Taylor's series expansions.

The rationale for f_k is quite simple. It is the expectation of a quadratic which coincides with u at the abscissa values $\mu - k\sigma$, μ, and $\mu + k\sigma$. It should, therefore, give a reasonable approximation when the shape of u is nearly quadratic.

An appealing feature of f_k is its simplicity. One need not calculate any derivatives. In fact, if u has derivatives of all orders, f_k may be written as

$$f_k(\mu,\sigma^2,u) = u(\mu) + \frac{\sigma^2}{2} u''(\mu) + \sum_{n=2}^{\infty} \frac{k^{2(n-1)}\sigma^{2n}u^{(2n)}(\mu)}{(2n)!} \qquad (3.5.5)$$

using the Taylor series expansion about μ. This representation is rather revealing in that it shows how f_k might adjust to non-quadratic u better than 3.5.2 by the inclusion of higher order derivatives. (Recall 3.5.2 is exact for quadratic u.)

What would be a reasonable way to evaluate how "good" f_k is? The results of Chapter two provide some insight here. For if $E_F[u]$ is known to lie in an interval [L, U], then a "reasonable" approximation to $E_F[u]$ should also be contained in that interval. Unfortunately, f_k does not possess this property. On balance, no approximator would be likely to be so contained unless it somehow took L and U into account.

To examine this issue of falling into (L, U) some plots will prove quite useful. Some popular utility functions scaled to [0, 1] will be considered. All possible μ, σ^2 pairs are contained in the region bounded by the x-axis and the function $g(x) = x(1-x)$, where the μ values are along the abscissa, the σ^2 values along the ordinate. Call this region Λ . The following two functions are plotted on that region.

Define

$$C_{Lk}^{u} = \{(\mu,\sigma^2)\,|\,\sigma^2 > 0,\ (\mu,\sigma^2) \in \Lambda,\ L(\mu,\sigma^2, u) = f_k(\mu,\sigma^2, u)\} \tag{3.5.6}$$

and

$$C_{Uk}^{u} = \{(\mu,\sigma^2)\,|\,\sigma^2 > 0,\ (\mu,\sigma^2) \in \Lambda,\ U(\mu,\sigma^2,u) = f_k(\mu,\sigma^2,u)\} \tag{3.5.7}$$

Note that C_{Lk}^{u} extends from (0, 0) upward to a point, say M_{ku} , on g . From there, C_{Uk}^{u} extends downward to (1, 0), although neither C_{Uk}^{u} nor C_{Lk}^{u} is necessarily monotone. Along g, $L = U = E_F[u]$ so that M_{ku}, $M_{ku}(1-M_{ku})$ is the μ,σ^2 pair ($\sigma^2 > 0$) for which $f_k(\mu,\sigma^2,u) = E_F[u]$. Above C_{Lk}^{u} the relation

$$f_k(\mu,\sigma^2,u) < L(\mu,\sigma^2,u) \tag{3.5.8}$$

holds. Above C_{Uk}^{u} the relation

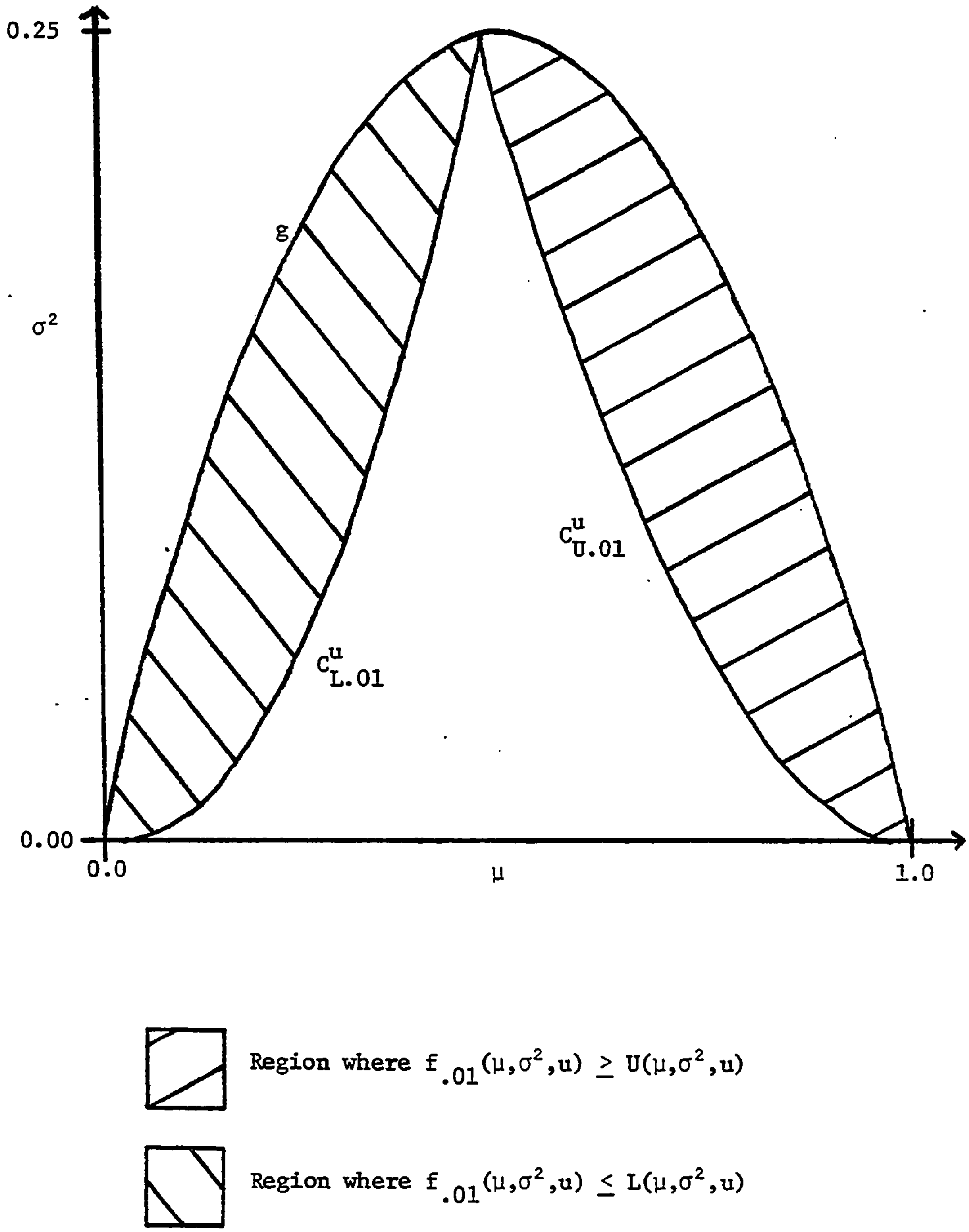
$$f_k(\mu,\sigma^2,u) > U(\mu,\sigma^2,u) \tag{3.5.9}$$

holds. As is evident from the graphs, these areas are distressingly large. The graphs are on the following pages. Note that $k = .01$ for all of these graphs.

These graphs do clearly illustrate the regions where f_k lies outside [L, U]. However, they do not show how that even when f_k lies in the interval, it may lie quite close to one end or the other. It should be recalled that it is the most extreme distribution that yields the bound. Using an f_k value near a bound is, perhaps, being either rather optimistic or pessimistic. This may be acceptable but before using f_k one should be aware of this behavior.

Before accepting this analysis, some questions regarding its validity must be answered. One might think it unfair to examine this approximator in the restricted sense of Chapter two, i.e., at least one finite endpoint and a certain "shape". However, it is clear that Levy and Markowitz had just such a setting in mind rather than a broader one. In their paper, it is postulated that the rate of return, R, is always > -1. In fact, they go further in the subsequent analysis to take R to always be in the interval $[-.3, .6]$.

The shapes they considered were also "correct" for the present methodology. Each utility function they used has $u' > 0$, $u'' < 0$, and $u''' \geq 0$ (i.e., having convex derivatives is the key here). Thus, it is clear that their approximator is intended for the same type of situation for which this mathematical programming approach yields



Region where $f_{.01}(\mu,\sigma^2,u) \geq U(\mu,\sigma^2,u)$

Region where $f_{.01}(\mu,\sigma^2,u) \leq L(\mu,\sigma^2,u)$

Figure 3.5.1. Plot for $u(x) = (1-\exp(-x))/(1-\exp(-1))$

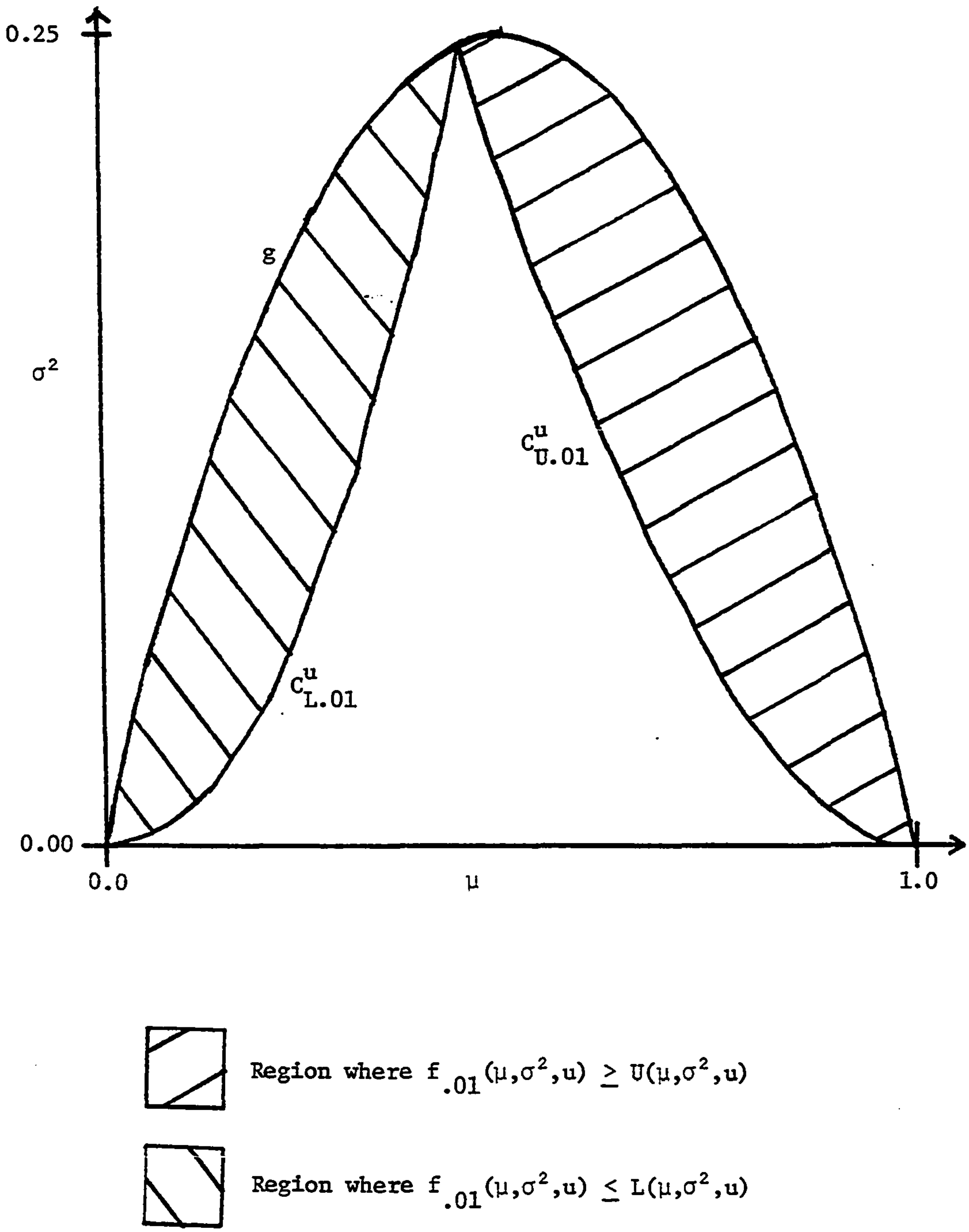
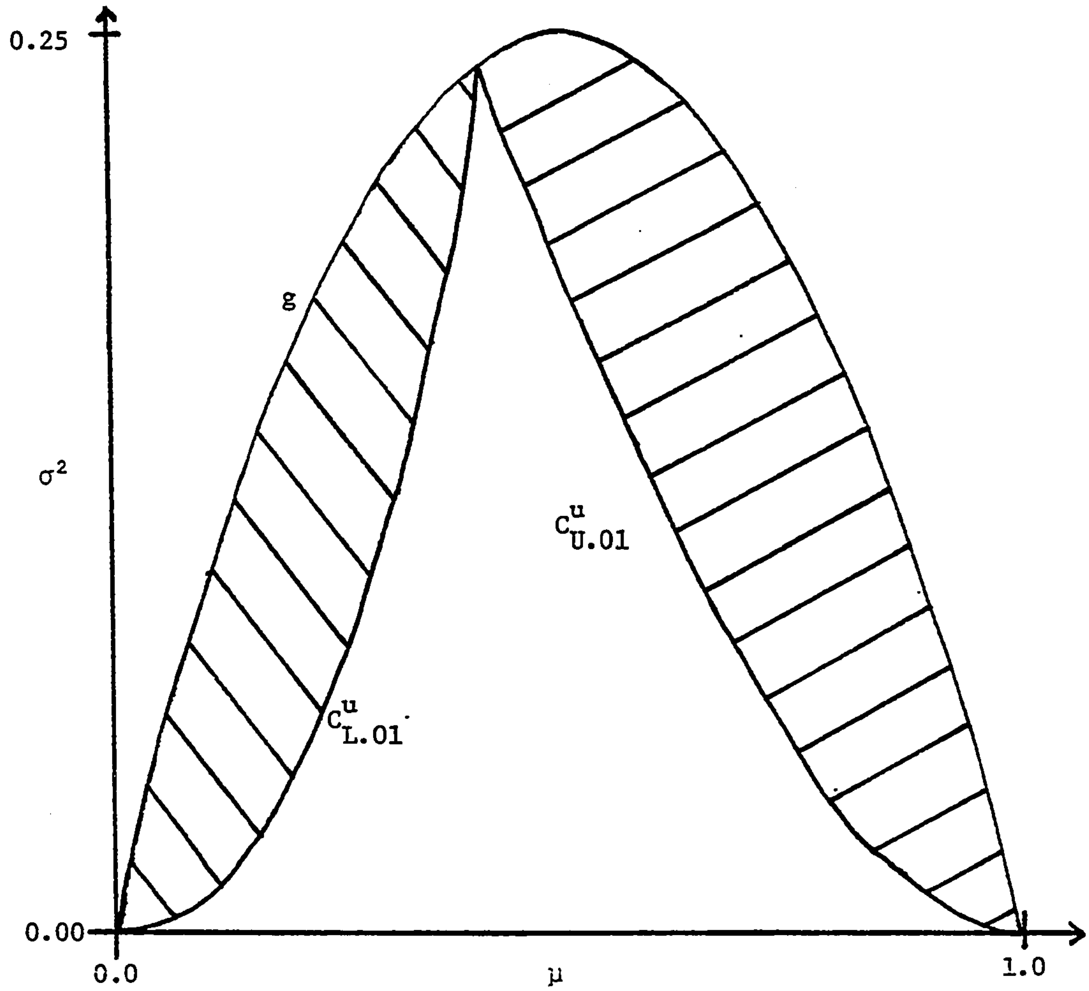


Figure 3.5.2. Plot for $u(x) = \ell n\ (1+x)/\ell n\ (2)$



Region where $f_{.01}(\mu,\sigma^2,u) \geq U(\mu,\sigma^2,u)$

Region where $f_{.01}(\mu,\sigma^2,u) \leq L(\mu,\sigma^2,u)$

Figure 3.5.3. Plot for $u(x) = (1-\exp(-3x))/(1-\exp(-3))$

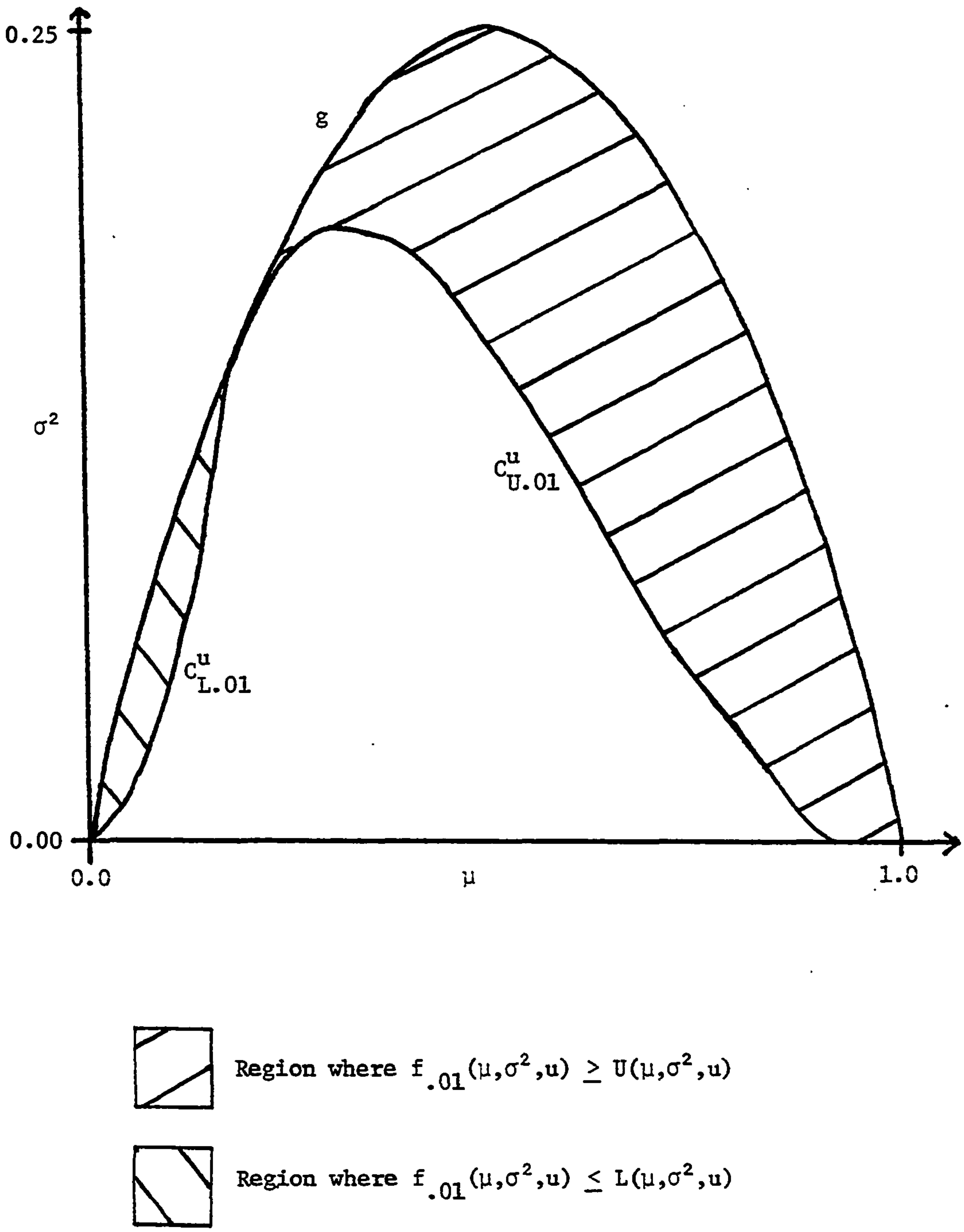


Figure 3.5.4. Plot for $u(x) = \sqrt{x}$

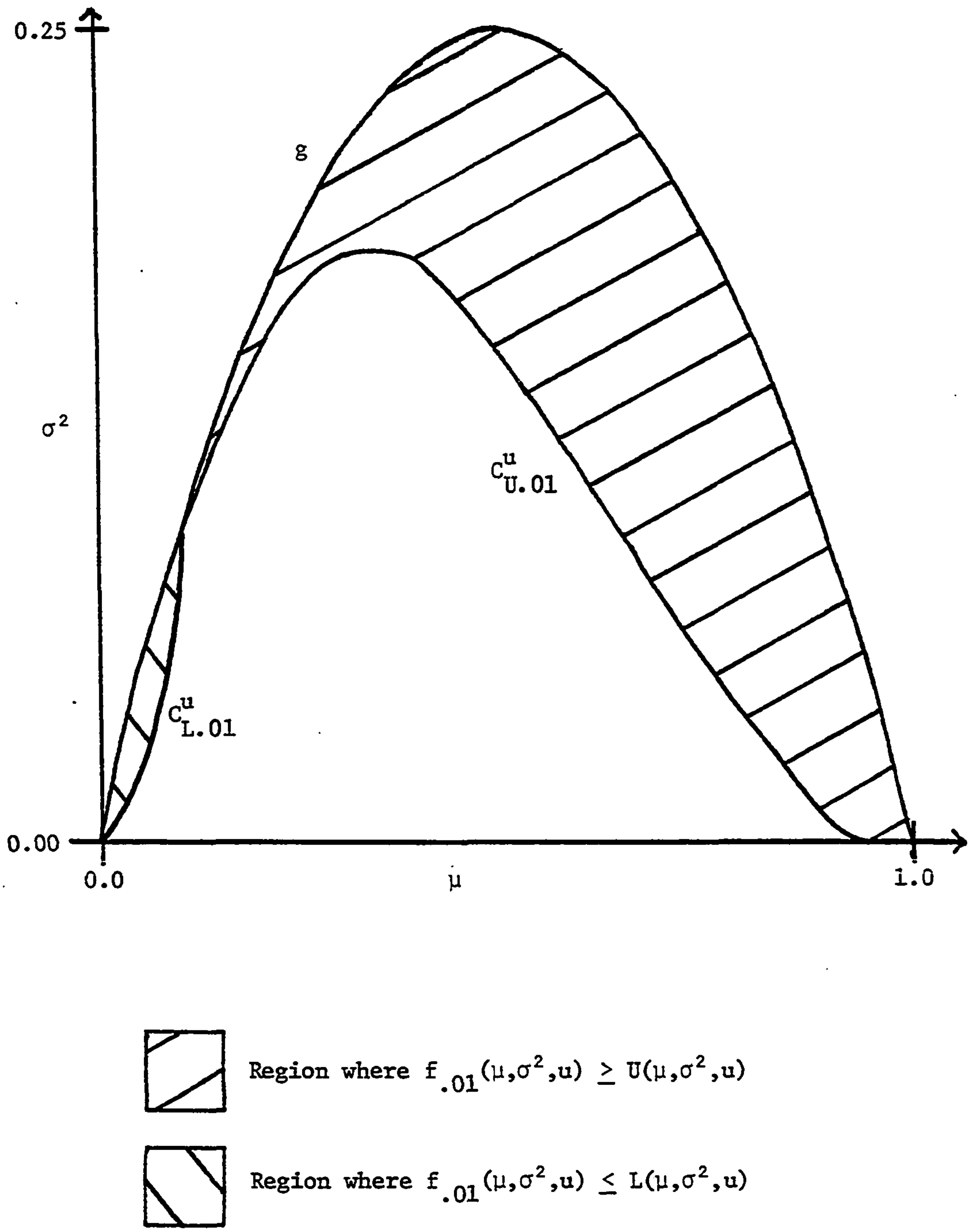


Figure 3.5.5. Plot for $u(x) = x^{1/4}$

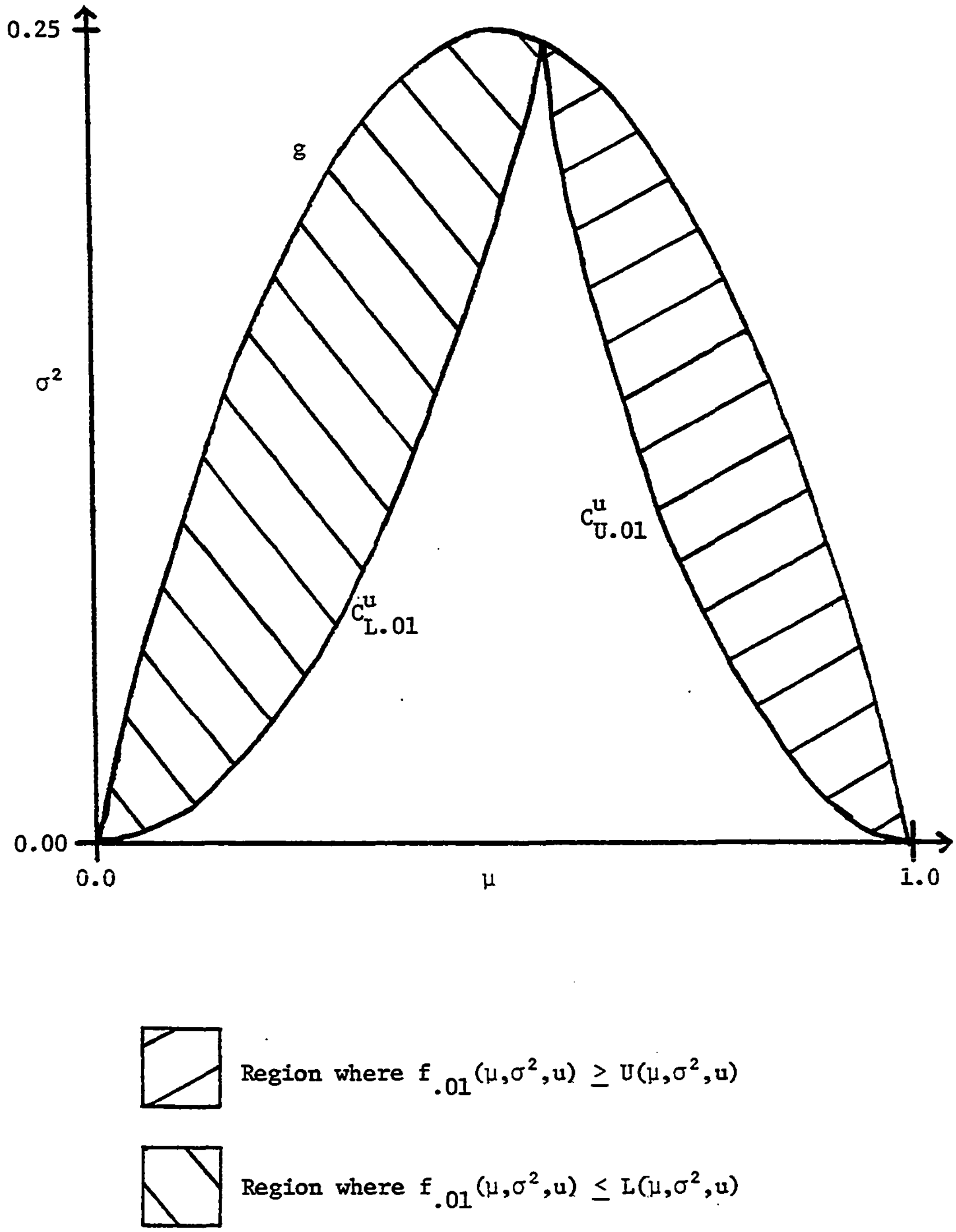


Figure 3.5.6. Plot for $u(x) = x^4$

bounds. It seems completely legitimate, therefore, to use those bounds to examine the behavior of f_k .

Another question concerns the choice of k . The value used for the graphs presented here is $k = .01$, the value suggested by Levy and Markowitz as being generally good. They do present suggestions for choosing a good value of k, and even indicate a case when .01 would be poor. However, the exponential curve presented in figure 3.5.1 is one for which they say $k = .01$ should be good. For the actual curves presented here, the choice of k seemed to make little difference. Several of the curves were examined for k in the range from .001 to 1 and almost identical results were obtained. It should be noted that determining C^u_{Lk} and C^u_{Uk} analytically is a formidable task. For producing the present plots, σ^2 values were determined by solving $f_k(\mu,\sigma^2,u) = L(\mu,\sigma^2,u)$ (for C^u_{Lk}, $U(\mu,\sigma^2,u)$ for C^u_{Uk}, of course) for selected values of μ . A numerical root finding technique, the regula falsi method, was employed to this end on an HP-41C programmable pocket calculator. See Berger and Hale (1980) for a discussion of that machine's usefulness in scientific applications. This proved reasonably efficient and is probably the simplest approach generally.

One might complain, however, about the finite interval. Suppose one has $[a, +\infty)$ or $(-\infty, a]$ for domain where a is finite. How does that affect this analysis? The answer is not particularly pleasing for users of the Levy-Markowitz approximator. This is because while g, the "lid", and one of the curves disappear, the other remains and, in fact, extends. The following figure illustrates this situation.

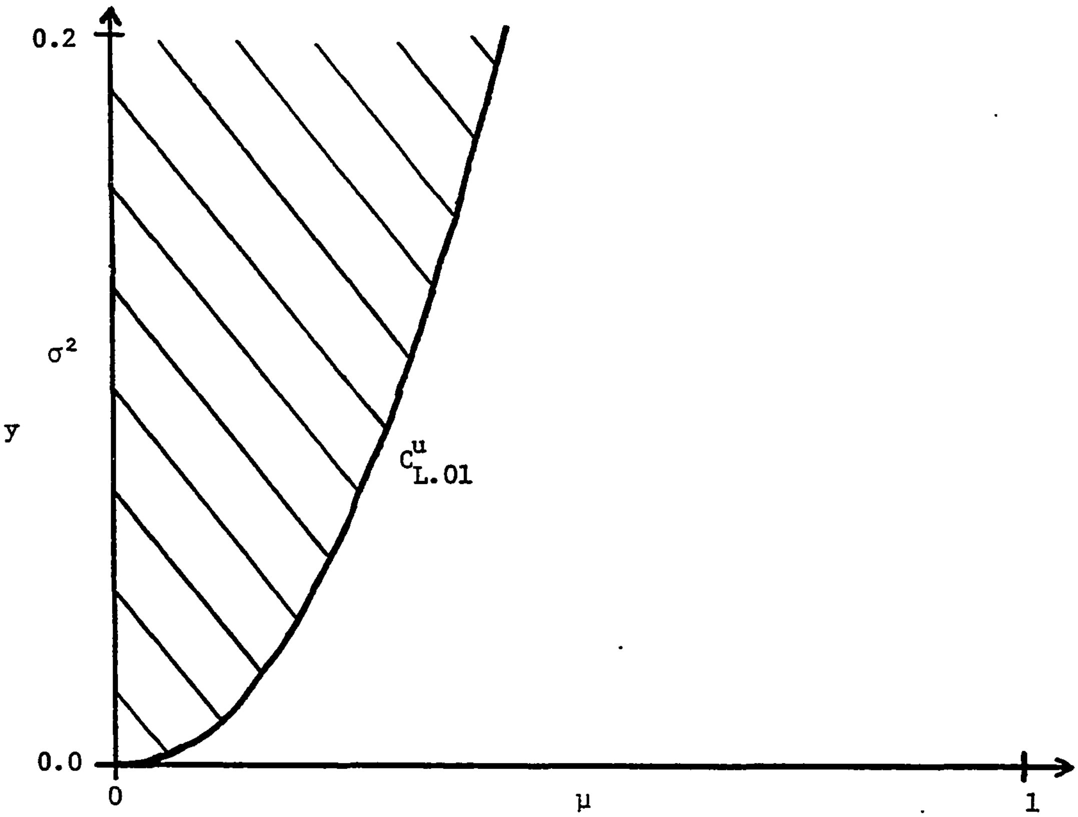


Figure 3.5.7. Plot for $u(x) = (1-\exp(-x))/(1-\exp(-1))$

Removing g gives many more σ^2 values which yield "bad" f_k approximations. Notice that in this figure the curve C^u_{Uk} no longer exists and hence, the region, which in figure 3.5.1 is the right shaded area, no longer produces values which are greater than U. Even if one were to use the Jensen upper bound, such a "bad" region cannot be produced. This is because for this function, f_k is always less than U for $\sigma^2 > 0$. Thus, while getting rid of one region, the other increases dramatically. Once again the importance of this exercise is in gaining an understanding of the behavior of f_k .

The results presented here are not intended to be exhaustive nor definitive. They do quite clearly point out one deficiency of the Levy-Markowitz approximator. One might try to adjust the approximator by "moving it back into" the interval when outside, but then why not simply take $(U + L)/2$? Further research is indicated here, particularly for the case when the distribution is known to be continuous.

It is interesting to note that part of this section could have been developed using Mantell's work. Mantell's (1976) paper is important in that it appears to be the first attempt to find a bound for expected utility. Finally, this section has made a start toward coordinating the theory of approximation and the theory of bounds. This worthwhile goal will aid the decision maker to remove part of the uncertainty inherit in the decision process.

IV. EXTENSIONS AND EXAMPLES

1. Introduction

This chapter is presented to supplement and illustrate ideas presented in previous chapters. In particular, much of chapter two could not profitably be elucidated until sufficient theory was developed. Theoretical treatment is touched only lightly in section three where two new results are presented, and also somewhat in section four where the general problem is considered and one new result proffered.

The aim throughout this chapter is to illuminate concepts and suggest various possibilities for extension. To accomplish this, most examples are not covered in depth. Proof is offered only when it seems most necessary. This chapter may appear rather episodic. The alternatives, including fewer examples and ideas or producing too voluminous a chapter, seem less desirable.

Section three clarifies several notions from chapter two. The fourth section discusses difficulties associated with a general solution for the problem $P_1: \min_{F \in \mathcal{F}} \int_a^b g\,dF$. Following that, specific applications to utility theory are given, in contrast with the general treatment in chapter three. A brief conclusion closes the chapter. The next section, section two, presents some rather explicit possibilities for application in reliability.

2. Applications in Reliability

Reliability is one possible field of application for this mathematical programming technique for bounding generalized moments. This is not surprising since physical realities often impose some natural constraints on the values which certain variables may take. This section will present some possibilities for consideration.

Example 4.2.1 A manufacturer produces parts having exponential life with density function

$$f_\lambda(x) = \lambda e^{-\lambda x} \, . \tag{4.2.1}$$

The parameter λ is well-known as the hazard rate or failure rate. (A common interpretation is that the probability of failure in a small interval of length Δt is approximately $\lambda\Delta t$, given that the unit is operational at the beginning of that interval.) Suppose that the parameter λ is a random variable. One way to view this is to suppose that the hazard rate is "built into" each part at the time of manufacture.

The manufacturer has reason to believe that $\lambda \in [a,b]$ with mean μ_1 and second moment μ_2. The programming approach of chapter two may be used to find bounds for the part population mean, $E[1/\lambda]$, and variance, $E[1/\lambda^2]$. (Actually, bounds for all moments $E[n!/\lambda^n]$ may be found using this approach. This is clear since $\frac{d}{d\lambda}(n!/\lambda^n)$ is concave on $[a,b]$ for all natural numbers n, $a > 0$.) One obtains

$$\frac{\frac{\sigma^2}{b} + \frac{(\mu_1-b)^3}{(\mu_2-b\mu_1)}}{\sigma^2 + (\mu_1-b)^2} \leq E[1/\lambda] \leq \frac{\frac{\sigma^2}{a} + \frac{(\mu_1-a)^3}{(\mu_2-a\mu_1)}}{\sigma^2 + (\mu_1-a)^2} \tag{4.2.2}$$

and

$$\frac{\frac{\sigma^2}{b^2} + \frac{(\mu_1-b)^4}{(\mu_2-b\mu_1)^2}}{\sigma^2 + (\mu_1-b)^2} \leq E[1/\lambda^2] \leq \frac{\frac{\sigma^2}{a^2} + \frac{(\mu_1-a)^4}{(\mu_2-a\mu_1)^2}}{\sigma^2 + (\mu_1-a)^2} \tag{4.2.3}$$

where $\sigma^2 = \mu_2 - \mu_1^2$.

To get an appreciation for these bounds, some actual numbers will now be used. Suppose that the hazard rate lies between .001 failures/hour and .02 failures/hour, i.e., $[a,b] = [.001, .02]$. Further, suppose that $\mu_1 = .01$ and $\mu_2 = .00010625$. It would, of course, be incorrect to state the mean part life as 100 hrs = 1/.01 . The bounds from 4.2.2 and 4.2.3 yield

$E[1/\lambda] \in [103,158]$

and

$E[1/\lambda^2] \in [10856, 79750]$.

For comparative purposes, if the second moment were smaller, say $\mu_2 = .000101$, then the intervals would be reduced, as follows:

$E[1/\lambda] \in [101, 110]$

and

$E[1/\lambda^2] \in [10127, 21857]$. $\square$

The next two examples are similar to that considered by Kim (1978). The essential point of difference is the restricted domain so that the results of chapter two may be applied.

Example 4.2.2 A new type of machine has a strength distribution which is known to the first two moments. The machine is subjected to a stress which has a known distribution. Let Y be the random variable denoting the stress, having cdf H which is completely known. Also, let X be the random variable for the strength of the machine. Suppose that X has cdf F, only known to the first two moments μ_1 and μ_2. This is not an entirely artificial setup. Indeed, the stress distribution may be fairly well-known after many years of experience whereas little may be known about the characteristics of a new machine.

Although Kim (1978) considered the problem of finding max $P(X>Y)$, in this example the problem of finding min $P(X>Y)$ is considered. This is useful since min $P(X>Y)$ can be thought of as a guaranteed survival probability.

Recall that

$$P(X>Y) = \int_0^{+\infty} H(t)dF(t) \ .$$

Hence, if any of the four conditions of Theorem 2.6.1 holds, then the attainable bound in 2.6.1 gives the desired minimum. For these examples only the simplest condition will be used. That is, two cases will be considered where the stress density function, H', is convex.

For the first case, consider a Weibull stress distribution, say $H(y) = 1 - \exp(-\lambda y^{\alpha})$. For what values of λ and α will the density function be convex? Let h be the density function ,

$$h(y) = \lambda\alpha y^{\alpha-1} \exp(-\lambda y^{\alpha}) \quad \lambda > 0 \ , \ \alpha > 0 \ , \ y \geq 0 \tag{4.2.4}$$

Then

$$\frac{dh}{dy} \propto \{(\alpha-1)y^{\alpha-2} - \lambda\alpha y^{2\alpha-2}\} \exp(-\lambda y^{\alpha}) \tag{4.2.5}$$

and

$$\frac{d^2h}{dy^2} \propto \lambda^2\alpha^2 y^{\alpha-3} \exp(-\lambda y^{\alpha}) \left\{ y^{2\alpha} - \frac{3(\alpha-1)}{\lambda\alpha} y^{\alpha} + \frac{(\alpha-1)(\alpha-2)}{\lambda^2\alpha^2} \right\} \tag{4.2.6}$$

For determining the sign of h'', only the part of 4.2.6 in brackets need be considered. This may be rewritten as

$$\left\{ y^{2\alpha} - \frac{3(\alpha-1)}{\lambda\alpha} y^{\alpha} + \frac{(\alpha-1)(\alpha-2)}{\lambda^2\alpha^2} \right\} = \left\{ y^{2\alpha} - 2 \left(\frac{3(\alpha-1)}{2\lambda\alpha}\right) y^{\alpha} + \frac{9(\alpha-1)^2}{4\lambda^2\alpha^2} \right\} + \frac{(\alpha-1)(\alpha-2)}{\lambda^2\alpha^2} - \frac{9(\alpha-1)^2}{4\lambda^2\alpha^2} \tag{4.2.7}$$

$$= [y^{\alpha} - \frac{3(\alpha-1)}{2\lambda\alpha}]^2 + \frac{(\alpha-1)(1-5\alpha)}{4\lambda^2\alpha^2} \ . \tag{4.2.8}$$

This is non-negative whenever $\alpha \varepsilon [1/5,1]$. Hence, for any value of $\lambda > 0$ and whenever $\alpha \varepsilon [1/5,1]$, h is convex. Thus, from 2.6.1

$$\min P(X>Y) = H(\mu_2/\mu_1)\mu_1^2/\mu_2 \tag{4.2.9}$$

$$= \frac{\mu_1^2}{\mu_2} \{1 - \exp(-\lambda (\frac{\mu_2}{\mu_1})^{\alpha})\} \ . \tag{4.2.10}$$

Some actual numbers will now be given for clarity. Suppose that $\alpha = 2/3$ and $\lambda = 2$. Thus, the stress has mean

$$\frac{\Gamma(1+1/\alpha)}{\lambda^{1/\alpha}} = \frac{\Gamma(5/2)}{2^{3/2}} = 0.47$$

and variance

$$\frac{\Gamma(1+2/\alpha) - \{\Gamma(1+1/\alpha)\}^2}{\lambda^{2/\alpha}} = .53 .$$

Further, suppose that the strength mean, μ_1, is 1.5 (roughly three times the stress mean) and that the strength variance is 0.5 (comparable to the stress variance). Then, using 4.2.10,

$$\min P(X>Y) = \frac{(1.5)^2}{2.75} \{1-\exp[-2(2.75/1.5)^{2/3}]\}$$

$$= .7773 .$$

Therefore, for the given assumptions the strength of the new machine will exceed the stress to which it is subjected with probability at least .7773 . $\square$

<u>Example 4.2.3</u> Suppose now that the stress has a gamma distribution, i.e., the density function is

$$h(y) = \frac{\lambda^{\alpha} y^{\alpha-1} \exp(-\lambda y)}{\Gamma(\alpha)} \quad \lambda > 0 , \alpha > 0 , y \geq 0 . \tag{4.2.11}$$

Hence,

$$\frac{dh}{dy} \propto \{(\alpha-1)y^{\alpha-2} - \lambda y^{\alpha-1}\} \exp(-\lambda y) \tag{4.2.12}$$

and

$$\frac{d^2h}{dy^2} \propto \lambda^2 y^{\alpha-3} \exp(-\lambda y)\{y^2 - \frac{2(\alpha-1)y}{\lambda} + \frac{(\alpha-1)(\alpha-2)}{\lambda^2}\} . \qquad (4.2.13)$$

Again, the sign of h'' is determined only by the bracketed portion. This is re-expressed as

$$\{y^2 - \frac{2(\alpha-1)}{\lambda} y + \frac{(\alpha-1)(\alpha-2)}{\lambda^2}\} =$$

$$\{y^2 - 2\left(\frac{\alpha-1}{\lambda}\right)y + \left(\frac{\alpha-1}{\lambda}\right)^2\} + \frac{(\alpha-1)(\alpha-2)}{\lambda^2} - \frac{(\alpha-1)^2}{\lambda^2} \qquad (4.2.14)$$

$$= [y - \left(\frac{\alpha-1}{\lambda}\right)]^2 + \frac{\alpha-1}{\lambda^2}[(\alpha-2) - (\alpha-1)] \qquad (4.2.15)$$

$$= [y - \left(\frac{\alpha-1}{\lambda}\right)]^2 + \frac{1-\alpha}{\lambda^2} . \qquad (4.2.16)$$

Thus, h is convex for all $\lambda > 0$, $\alpha \varepsilon (0,1]$.

As previously, from 2.6.1

$$\min P(X>Y) = H(\mu_2/\mu_1)\mu_1^2/\mu_2 .$$

To illustrate, suppose that $\alpha = .4168$ and $\lambda = .8868$, so that the stress mean is $\frac{\alpha}{\lambda} = .47$ and the variance is $\frac{\alpha}{\lambda^2} = .53$. Now suppose that the strength mean is 1.0 and the strength variance is 0.833. Then

$$\min P(X>Y) = \frac{(1)^2}{(1.833)} \int_0^{1.833} \frac{(.8868)^{.4168} y^{-.5832} \exp(-.8868y)}{\Gamma(.4168)} dy$$

$$= 0.5198 ;$$

i.e., the strength of the machine will exceed the stress with a probability of at least 0.5198 . □

This section has presented a brief indication of some possibilities for applying mathematical programming to reliability. In the next section, additional insight is provided into issues raised in chapter two.

3. Additional Extensions and Interrelations

Many ideas and results were presented in chapter two. So many, in fact, that several things were deferred to this chapter where a more informal approach proves expeditious. The purpose of this section is to illuminate some of those ideas through illustration. Figure 4.3.1 is also provided, showing the relationships among the main conditions of chapter two. Additionally, two new results are presented. The first of these constitutes Example 4.3.2 . The other is presented in Example 4.3.4.

The first example shows how Theorem 2.6.1 may be applied for bounding moments.

<u>Example 4.3.1</u> Consider the non-negative random variable X with known first and second moments, μ_1 and μ_2 respectively. Define

$$\mu_\alpha = E[X^\alpha]$$

$$= \int_0^{+\infty} x^\alpha dF(x) \tag{4.3.1}$$

where $\alpha > 0$ and F is the distribution function of X . Notice that

$$\frac{d}{dx}(x^\alpha) = \alpha x^{\alpha-1}$$

is strictly convex for $\alpha \in (2,+\infty)$ or $\alpha \in (0,1)$, strictly concave for $\alpha \in (1,2)$, and uninteresting for $\alpha = 2$ or $\alpha = 1$. Hence, appealing to Theorem 2.6.1,

$$\frac{\mu_2^{\alpha-1}}{\mu_1^{\alpha-2}} \leq \mu_\alpha \text{ for } \alpha \in (0,1),\ \alpha \in (2,+\infty) \tag{4.3.2}$$

and

$$\mu_\alpha \leq \frac{\mu_2^{\alpha-1}}{\mu_1^{\alpha-2}} \text{ for } \alpha \in (1,2) \quad . \tag{4.3.3}$$

Note that 4.3.2 is a fairly well-known inequality for α a natural number greater than two.

Another pair of bounds is obtained if X is restricted to lie in a finite interval. Take the interval to be [0,1] since a simple transformation (division by β when X lies in $[0,\beta]$) will always permit this. Combining the new bounds with the ones above yields the following inequalities

$$\frac{\mu_2^{\alpha-1}}{\mu_1^{\alpha-2}} \leq \mu_2 \leq \frac{\sigma^2 + (\frac{\mu_2-\mu_1}{\mu_1-1})^\alpha (\mu_1-1)^2}{(\mu_1-1)^2 + \sigma^2}$$
$$\text{for } \alpha \in (0,1) \ , \ \alpha \in (2,+\infty) \tag{4.3.4}$$

and

$$\frac{\sigma^2 + (\frac{\mu_2-\mu_1}{\mu_1-1})^\alpha (\mu_1-1)^2}{(\mu_1-1)^2 + \sigma^2} \leq \mu_\alpha \leq \frac{\mu_2^{\alpha-1}}{\mu_1^{\alpha-2}} \text{ for } \alpha \in (1,2) \tag{4.3.5}$$

where $\sigma^2 = \mu_2 - \mu_1^2$ as before. (Note that, if rescaling is done, μ_1 and μ_2 must be rescaled accordingly.) $\square$

It is quite clear from this that all moments of a finite, non-negative random variable are constrained by the first two moments to lie in precisely defined intervals. As noted several times previously in this work, those intervals cannot be improved without adding further assumptions concerning the distribution of that random variable. Further work might examine the effect of symmetry or unimodality on the bounds. Additional assumptions might include the existence of a density function, monoticity of that function, or some other shape constraint.

The next example will present conditions for

$$\min_{F\varepsilon\mathcal{F}} \int_a^b (\sum_I w_i g_i) dF = \sum_I w_i \min_{F\varepsilon\mathcal{F}} \int_a^b g_i dF$$

to hold. This result is actually quite trivial. It is presented here more for the sake of illustration than for its theoretical importance. It underscores the prominence of objective function shape in this programming theory.

<u>Example 4.3.2</u> Consider the problem

$$\min_{F\varepsilon\mathcal{F}} \int_a^b (\sum_I w_i g_i) dF \tag{4.3.6}$$

where

1. $\mathcal{F} = \{F|F$ is a cdf on $[a,b]$ with $\int_a^b x dF = \mu_1$, $\int_a^b x^2 dF = \mu_2\}$,
2. I is an index set,
3. g_i is continuous at a and has continuous derivative on (a,b), $\forall\ i\ \varepsilon\ I$,

4. $g_i' \geq Q_t'$ on $[t,b)$, $\forall\, t \in (a,b)$, $\forall\, i \in I$,

5. $w_i \geq 0$, $\forall\, i \in I$,

   and

6. $\sum_I w_i = 1$.

Under these conditions one has

$$g' \geq Q_t' \text{ on } [t,b) \ \forall\, t \in (a,b) \tag{4.3.7}$$

where $g = \sum_I w_i g_i$. This is because

$$g_i' \geq Q_t' \text{ on } [t,b) \ \forall\, t \in (a,b),\ \forall\, i \in I$$

$$\Rightarrow \quad w_i g_i' \geq w_i Q_t' \text{ on } [t,b) \ \forall\, t \in (a,b),\ \forall\, i \in I$$

$$\Rightarrow \quad g' = \sum_I w_i g_i' \geq \sum_I w_i Q_t' = Q_t' \text{ on } [t,b),\ \forall\, t \in (a,b) \ . \tag{4.3.8}$$

From Theorem 2.6.1

$$\frac{g(a)(\mu_2-\mu_1^2) + g\left(\frac{\mu_2-a\mu_1}{\mu_1-a}\right)(\mu_1-a)^2}{(\mu_1-a)^2 + (\mu_2-\mu_1^2)}$$

is the unique lower attainable bound for 4.3.6 . This may be re-expressed as

$$\frac{(\mu_2-\mu_1^2)\ \sum_I w_i g_i(a) + (\mu_1-a)^2\ \sum_I w_i g_i\left(\frac{\mu_2-a\mu_1}{\mu_1-a}\right)}{(\mu_1-a)^2 + (\mu_2-\mu_1^2)}$$

which simplifies to

$$\sum_I w_i \left\{ \frac{g_i(a)(\mu_2-\mu_1^2) + g(\frac{\mu_2-a\mu_1}{\mu_1-a})(\mu_1-a)}{(\mu_1-a)^2 + (\mu_1-\mu_1^2)} \right\}$$

$$= \sum_I w_i \min_{F\varepsilon\mathcal{F}} \int_a^b g_i dF \quad .$$

Hence,

$$\min_{F\varepsilon\mathcal{F}} \int_a^b (\sum_I w_i g_i) dF = \sum_I w_i \min_{F\varepsilon\mathcal{F}} \int_a^b g_i dF \quad . \qquad \square \qquad (4.3.9)$$

It should be noted that this result could well be included with those in section six of chapter two. That is, one way to show that $g \geq Q_t$ is to re-express g as $\sum_I w_i g_i$ satisfying conditions 2 through 6 above. Another way to look at it is that weighted sums of "feasible functions" are themselves feasible.

The remainder of this section will demonstrate how the various main conditions of chapter two are related. Figure 4.3.1 illustrates the relationships among the following conditions.

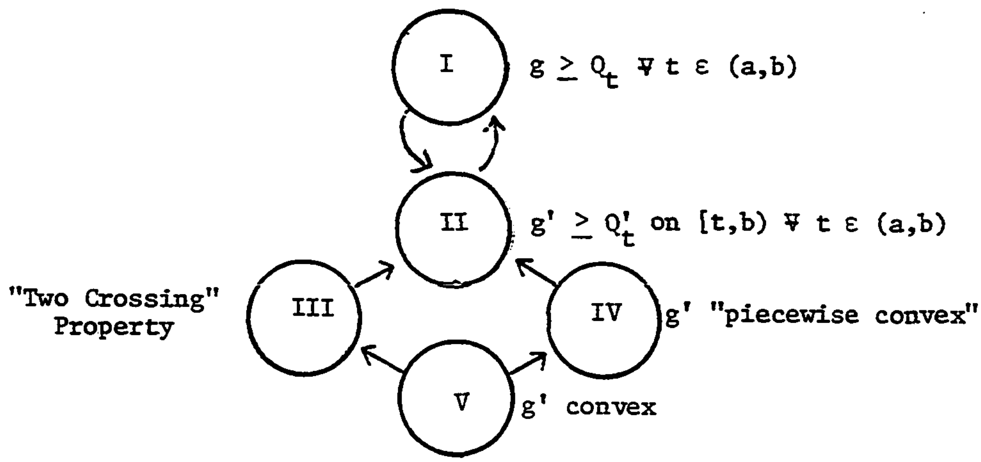


Figure 4.3.1. Relationship among conditions

I. $g \geq Q_t$ on $[a,b]$, $\forall\, t \in (a,b)$. (Theorem 2.4.1).

II. $g' \geq Q_t'$ on $[t,b)$, $\forall\, t \in (a,b)$. (Theorem 2.4.1).

III. g' coincides with Q_t' on at most two distinct intervals with $Q_t'(a) < g'(a)$ for some t (unless $g \equiv Q_t$) . (Theorem 2.5.2).

IV. $\exists\, \{a_i\}_{i\in I} \ni g' \geq Q_t'$ on $[t,a_1]$ $\forall\, t \in (a,a_1)$, g' is convex on $[a_i,a_{i+1})$, $\forall\, i \in I$, and $g' \geq Q_{a_i}'$ on $[a_i,a_{i+1})$, $\forall\, i \in I$ (Corollary 2.5.4).

V. g' is convex (Corollary 2.5.2).

The following example demonstrates that IV implies neither III nor V and that II (and hence, I) implies none of III, IV, and V.

Example 4.3.3 Let $[a,b]$ be $[0,5]$, $g(0) = 0$, and

$$g'(x) = \begin{cases} 16 - 16x & 0 \leq x \leq 1 \\ 3(x-1)^2 & 1 < x \leq 2 \\ 6 - (3/2)x & 2 < x \leq 4 \\ 3(x-4)^2 & 4 < x \leq 5 \end{cases}$$

The curve g' is pictured in Figure 4.3.2. Clearly g' is not convex. Further, let $t_0 = 4\cos\theta + 2$ where $\theta = \frac{1}{3}\cos^{-1}(-5/8)$ (this is obtained from 2.5.15 by solving $\lambda_2 = 0$ for t). Hence, $t_0 \approx 4.93$ which yields $\lambda_2 = 0$ so that $Q_{t_0}' \doteq 2.6$. This Q_{t_0}' intersects g' distinctly at $x \doteq 0.84, 1.93, 2.27,$ and 4.93 . Thus, IV implies neither III nor IV.

Note that this function is one for which the weak duality approach of chapter two will not yield an upper bound. This is clear upon considering W_2^1, since $g' \not\geq W_2$ on $[0,2]$. (See Theorem 2.4.4.).

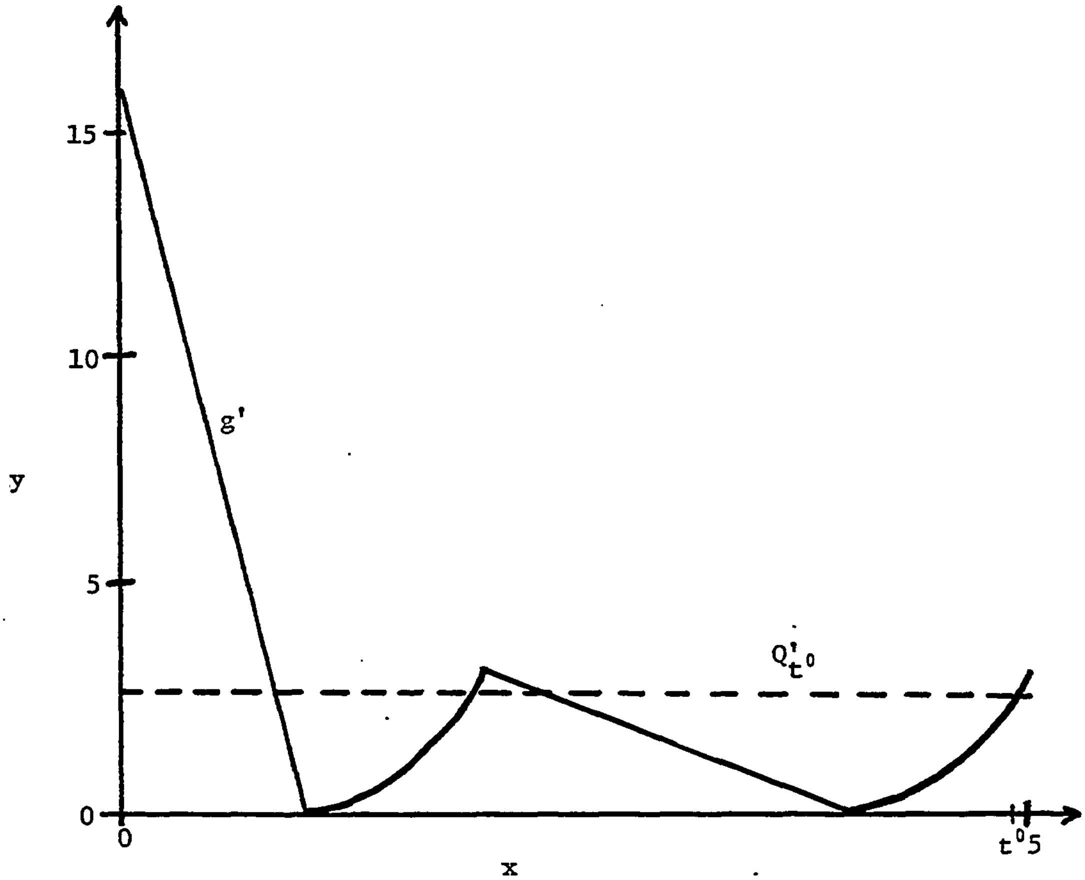


Figure 4.3.2. Derivative for Example 4.3.3

Now redefine $g'(x) = 1 - (x-5)^2$ on $(4,5]$. See Figure 4.3.3. Condition IV is now no longer satisfied, though condition II still holds. This is because $g' \geq Q'_t$ on $[t,5)$, $\forall\ t \in (0,4]$, g' is monotone increasing on $[4,5]$, and the slope of Q'_5 is $-46/75$, i.e., is negative. Theorem 4.3.1 explains why these conditions ensure condition II. However, it is clear that none of the conditions III, IV, or V are met. $\square$

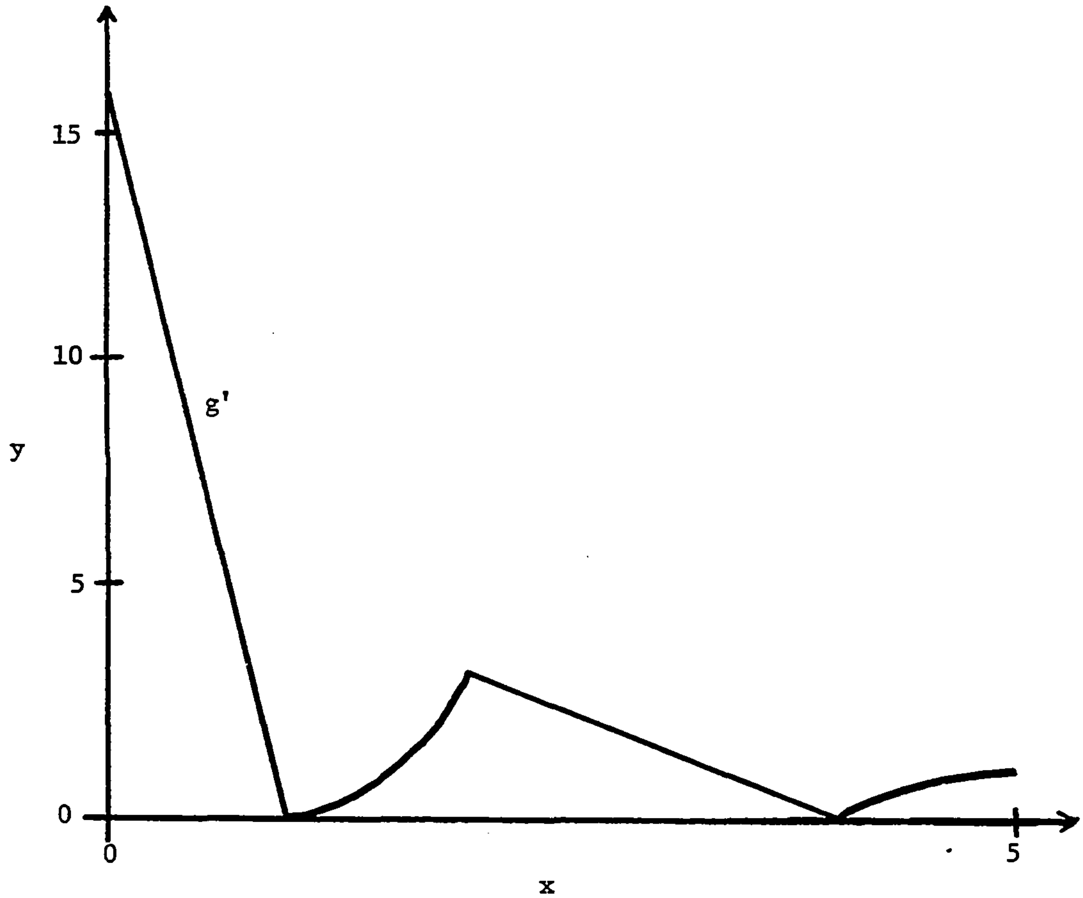


Figure 4.3.3. Modified derivative for Example 4.3.3

The next example shows that III implies neither IV nor V. Furthermore, a new result guaranteeing feasibility is presented.

<u>Example 4.3.4</u> Let [a,b] be [0,2], g(0) = 0, and

$$g'(x) = \begin{cases} 16 - 16x & 0 \leq x \leq 1 \\ 3 - 3(x-2)^2 & 1 < x \leq 2 \end{cases}$$

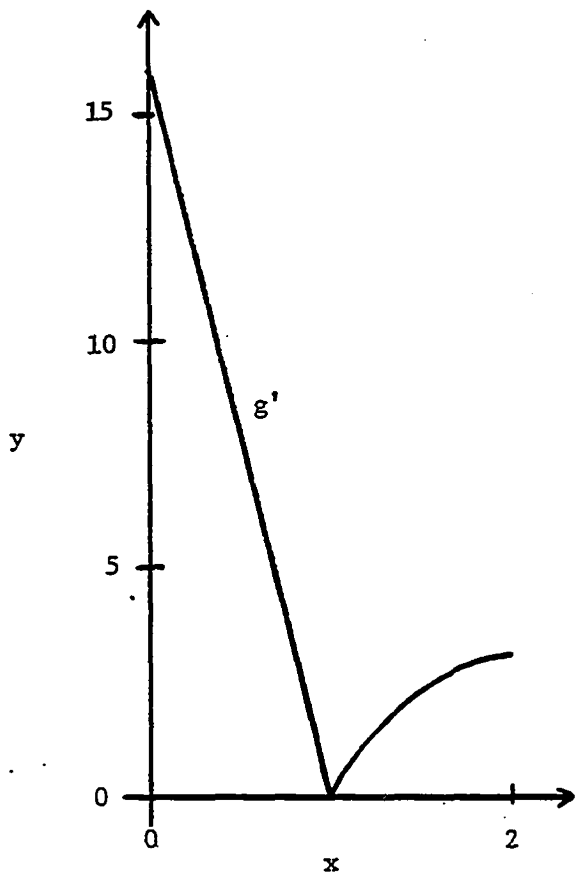


Figure 4.3.4. Derivative for Example 4.3.4

See Figure 4.3.4 for a depiction of g' . This curve satisfies condition III, but clearly neither IV nor V. It should be pointed out that this function g' satisfies a sufficient condition (which was not presented in chapter two) for $g \geq Q_t$. That condition is contained in

<u>Theorem 4.3.1</u>

Let g be a real valued function on $[a,b]$, continuous at a, with continuous derivative on (a,b). If there exists a point $c \varepsilon (a,b)$ such that

1. $g' \geq Q_t'$ on $[t,b)$, $\forall\ t \varepsilon (a,c]$, (4.3.10)

2. g' is monotone non-decreasing on $[c,b]$, (4.3.11)

and 3. the slope of Q_b' (i.e., $2\lambda_2$) is ≤ 0 (4.3.12)

then $g \geq Q_t$ on $[a,b]$, $\forall\ t \varepsilon (a,b)$.

Note that the conditions imply none of III, IV, or V and are implied by none of those same conditions. This result could well have been included in chapter two, but conditions 4.3.11 and 4.3.12 are of a somewhat more restrictive nature.

Theorem 4.3.1 is actually quite simple to show because the second and third conditions guarantee

$$g' \geq Q_t' \text{ on } [t,b)\ \forall\ t \varepsilon (c,b) . \tag{4.3.13}$$

That is, if 4.3.13 were not true, then there would exist a point $t_0 \varepsilon (c,b)$ such that Q_{t_0}' has positive slope. This is because g' is monotone non-decreasing on $[c,b]$.

However,

$$g'(b) \geq g'(t_0) \tag{4.3.14}$$

and

$$\text{the slope of } Q_b' \leq 0 , \tag{4.3.15}$$

so that

$$Q'_b > Q'_{t_0} \text{ on } [a,t_0] \,. \tag{4.3.16}$$

Furthermore,

$$Q'_b \geq g' \text{ on } [c,b] \,, \tag{4.3.17}$$

again by monotonicity of g' and the sign of the slope of Q'_b . Hence,

$$g(b) - g(a) = Q_b(b) - Q_b(a) \tag{4.3.18}$$

$$= \int_a^b Q'_b(x)dx$$

$$> \int_a^{t_0} Q'_{t_0}(x)dx + \int_{t_0}^b g'(x)dx \tag{4.3.19}$$

$$= \{g(t_0) - g(a)\} + \{g(b) - g(t_0)\}$$

$$= g(b) - g(a) \,, \tag{4.3.20}$$

a clear contradiction. Of course the result follows immediately from Theorem 2.4.1 . □

By this time the reader may have the impression that it is always necessary to show $g \geq Q_t$. Recall that Q_t is a special quadratic, however, always passing through $(a,g(a))$. The next section discusses the problem for more general g . A method is also presented which will generate a quadratic for any μ_1, μ_2 pair whenever g satisfies a certain set of conditinns. Those conditions will be seen to be quite different from those of chapter two. In fact, no g can satisfy both sets of conditions. The quadratic generated will also be of a special nature, and quite different from Q_t .

4. Concerning the General Solution

In this section the nature of a general solution for P_1 is considered. Recall that

$$P_1: \min_{F \in \mathcal{F}} \int_a^b g\,dF$$

where $\mathcal{F} = \{F \mid F \text{ is a cdf on } [a,b] \text{ with } \int_a^b x\,dF = \mu_1,\ \int_a^b x^2 dF = \mu_2\}$. A set of sufficient conditions is also presented which yields a solution to P_1 . In contrast to the conditions of the previous section, these do not fit into the general scheme of chapter two. Indeed, when they are satisfied, the condition $g' \geq Q_t'$ on $[t,b)$ $\forall$ $t \in (a,b)$ will not hold, except in a certain pathological case. Those conditions are given in 4.4.7-4.4.8.

It has been noted several times in the present work that, for g an arbitrary finite-valued function, a solution for P_1 is obtained by ad hoc means. This may appear to be at odds with the view point expressed by Karlin and Studden (1966), in their Theorem 2.1 (Chapter XII of Karlin and Studden (1966)), which proves the existence of a solution via weak duality. (Note: They never explicitly mention weak duality, though they demonstrate feasibility and equality of a proposed dual (though not called that) solution.) It will be remembered that P_1 has the weakly dual problem

$$D_1: \max_{(\lambda_0,\lambda_1,\lambda_2) \in \Lambda} \lambda_0 + \lambda_1\mu_1 + \lambda_2\mu_2$$

where $\Lambda = \{(\lambda_0,\lambda_1,\lambda_2) \mid Q(x) = \lambda_0 + \lambda_1 x + \lambda_2 x^2 \leq g(x)\ \forall\ x \in [a,b],$
$(\lambda_0,\lambda_1,\lambda_2) \in E^3\}$.

Theorem 2.1 is indeed a valuable theorem. However, finding an optimal $(\lambda_0,\lambda_1,\lambda_2)$ triplet may be an arduous task, even though Theorem 2.1 guarantees existence. Thus, at a certain level of generality a method of solution, i.e., weak duality, is guaranteed to be available, although actually determining $(\lambda_0,\lambda_1,\lambda_2)$ is still an ad hoc procedure. Some procedures have been suggested for solution when g is subject to certain conditions. This present work is one example. Another example is given in Kim (1979) where g is taken to be a cdf that is strictly convex on $(-\infty,0]$ and strictly concave on $[0,+\infty)$ with $\mu_1 = 0$. As pointed out in chapter one of the present work, he then solves

$$ps + (1-p)t = \mu_1 \tag{4.4.1}$$

$$ps^2 + (1-p)t^2 = \mu_2 \tag{4.4.2}$$

$$\lambda_0 + \lambda_1 s + \lambda_2 s^2 = g(s) \tag{4.4.3}$$

$$\lambda_0 + \lambda_1 t + \lambda_2 t^2 = g(t) \tag{4.4.4}$$

$$\lambda_1 + 2\lambda_2 s = g'(s) \tag{4.4.5}$$

$$\lambda_1 + 2\lambda_2 t = g'(t) \tag{4.4.6}$$

for $\lambda_0,\lambda_1,\lambda_2$, p, s, and t where $\lambda_2 > 0$, $t > s$, $\mu_1 = 0$, and $\mu_2 = \sigma^2$. The solution of equations 4.4.1-4.4.6 can also be employed for other functions g which are not of a convex-concave shape. The following two conditions ensure that 4.4.1-4.4.6 yields a feasible, and hence optimal, pair Q^*, F^* when s and t are such that $s + t = a + b$ and $s,t \in [a,b]$. The conditions for g a finite real-valued function with continuous derivative are

$$1) \quad g'(\tfrac{a+b}{2} + w) + g'(\tfrac{a+b}{2} - w) = 2g'(\tfrac{a+b}{2}) \ \forall \ w \in [0,\tfrac{b-a}{2}] \qquad (4.4.7)$$

and $2)$ g' is concave on $[a,\frac{a+b}{2}]$. (4.4.8)

Without loss of generality $(a+b)/2$ may be taken as zero. Then a solution for 4.4.1-4.4.6 is given by $s = \sqrt{\mu_2}$ (recall that $\mu_2 \leq b^2$) , $t = -\sqrt{\mu_2}$, $\lambda_0 = g(\sqrt{\mu_2}) - \frac{\sqrt{\mu_2}}{2}(g'(0) + g'(\sqrt{\mu_2}))$, $\lambda_1 = g'(0)$, and $\lambda_2 = (g'(\sqrt{\mu_2}) - g'(-\sqrt{\mu_2}))/2\sqrt{\mu_2}$.

While often useful, the application of 4.4.1-4.4.6 is not a panacea. The following example is intended to illustrate some of the problems. Possible remedies are also discussed.

Example 4.3.1 Consider the function

$g(x) = 1 - \cos(x)$ on $[0,2\pi]$. Note that g' does not satisfy $g' \geq Q_t'$ on $[t,2\pi)$ $\forall$ $t \in (0,2\pi)$. This means that the solution of chapter two is not valid for arbitrary (μ_1,μ_2) . The method cited above does work, however, since 4.4.7-4.4.8 are satisfied. This is because, as will be recalled, to solve P_1 through weak duality it is necessary to find a quadratic Q^* and a distribution F^* such that

$$\int_0^{2\pi} (g-Q^*)dF^* = 0 \qquad (4.4.9)$$

and

$$Q^* \leq g \text{ on } [0,2\pi] . \qquad (4.4.10)$$

When the conditions in 4.4.7-4.4.8 hold, a solution to 4.4.1-4.4.6 yields Q^* and F^* satisfying 4.4.9 and 4.4.10 .

To illustrate the kinds of difficulties which can arise when different μ_1,μ_2 pairs are evaluated, consider $(\mu_1,\mu_2) = (\pi,5\pi^2/4) = (\pi,12.337)$, and also $(\mu_1,\mu_2) = (\pi,15.915)$.

For the first moment pair, solution of 4.4.1-4.4.6 with the conditions $s + t = 2\pi$ and $s,t \in [0,2\pi]$ yields the optimal quadratic

$$Q^*(x) = -\frac{1}{\pi}x^2 + 2x + 1 - 3\pi/4$$

and the optimal distribution F^* which places mass 1/2 at $x = \pi/2$ and $x = 3\pi/2$. The lower bound is $1/2\{1 - \cos(\pi/2) + 1 - \cos(3\pi/2)\} = 1$.

For the second moment pair, solution of 4.4.1-4.4.6 with the conditions $s + t = 2\pi$ and $s,t \in [0,2\pi]$ yields the quadratic

$$W^*(x) = -0.1283\,x^2 + 0.8063\,x - 0.2665$$

and the distribution H^* which places mass 1/2 at $y = 0.6829$ and $y = 5.6003$. The lower bound is $1/2\{1 - \cos(0.6829) + 1 - \cos(5.6003)\} = 0.2242$. Figure 4.4.1 pictures g, Q^*, and H^* .

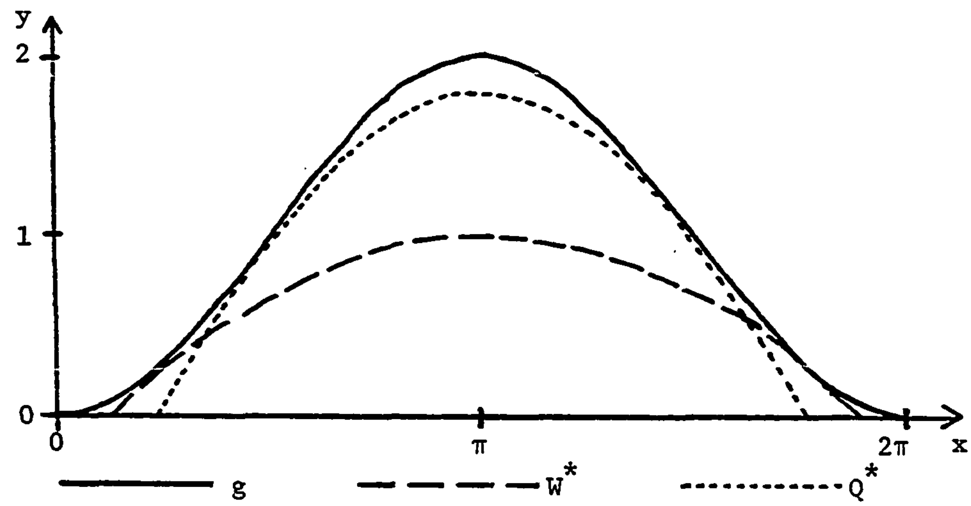


Figure 4.4.1. Cosine example

What happens now if g is changed slightly, as in Figure 4.4.2? For example, suppose now that one defines

$$g_{\delta,d}(x) = \begin{cases} g(x) & x \in [0,\pi-\delta] \cup [\pi+\delta,2\pi] \\ \frac{d-g(\pi-\delta)}{\delta}(x-\pi) + d & x \in (\pi-\delta,\pi] \\ \frac{g(\pi+\delta)-d}{\delta}(x-\pi) + d & x \in (\pi,\pi+\delta) \ , \end{cases}$$

$$\delta \in (0,\pi) \ , \ d \in [0,g(\pi-\delta)) \ ,$$

and let $g_{\delta,d}$ be the g of P_1 . The figure illustrates $\delta = 0.1$ and $d = 1$. As can be plainly seen, Q^* is no longer feasible, although W^* remains feasible . This illustrates that solving 4.4.1-4.4.6 does not necessarily produce a feasible solution. For the modification $g_{\delta,d}$ and the first moment pair, one needs to adjoin the

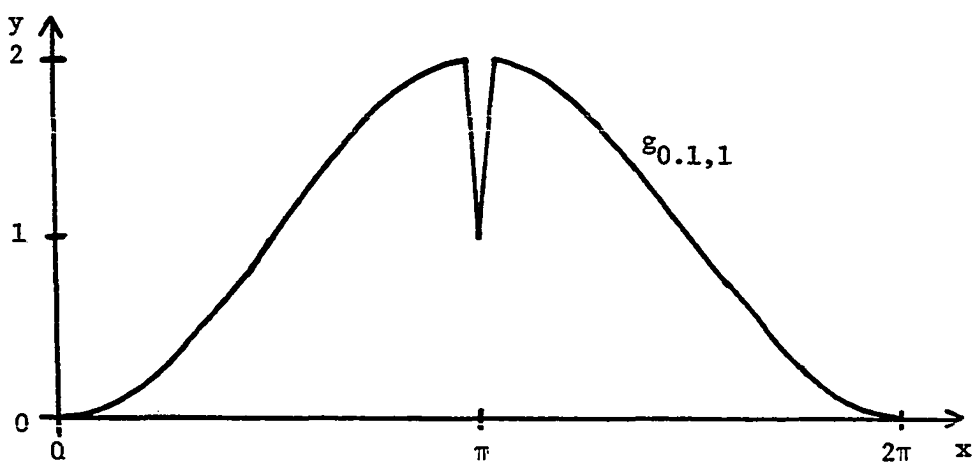


Figure 4.4.2. Modified cosine example

additional constraint $\lambda_0 + \lambda_1\pi + \lambda_2\pi^2 = d$, modify 4.4.1 and 4.4.2 to include the point $x = \pi$ as an atom of F^* , and therefore solve

$$\kappa(x-\pi)^2 + d = 1 - \cos x \tag{4.4.11}$$

$$2\kappa(x-\pi) = \sin x \tag{4.4.12}$$

for κ and x , $x \in [0,\pi]$. The left hand side of 4.4.11 is the optimal quadratic. The corresponding optimal distribution places mass $p = \pi^2/8(\pi-x)^2$ at x and $2\pi-x$, and mass $1-2p$ at π . In particular, for $d = 1$ as in the figure, the new optimal quadratic is

$$Q^*(x) = -\ 0.1283x^2 + 0.8063x - 0.2665 \ .$$

The corresponding optimal F^* places mass 0.2041 at $x = 0.6829$, mass 0.5918 at $x = \pi$, and mass 0.2041 at $x = 5.6003$. The desired lower bound is now

$$-\ 0.1283(5\pi^2/4) + 0.8063\pi - 0.2665 = 0.6834 \ .$$

Notice that for this example a three point distribution is necessary since a quadratic cannot "fit" on either side of the spike and "match" a distribution with mean π . Up to this point, only two point distributions have been considered.

Some other points should also be noted at this time. First, the small change from g to $g_{\delta,d}$ invalidated Q^* for the first moment pair but did not do so for the second moment pair. This is quite an undesirable feature for a method, in contrast to the assured feasibility, for any (μ_1,μ_2) , in the case of the weak duality approach of chapter two. Conditions 4.4.7–4.4.8 are helpful in guaranteeing feasibility of a solution to 4.4.1–4.4.6, but a set of necessary and sufficient conditions would be desirable for the method to be readily applicable.

Another point to note is that the new Q^* equals H^*. In fact, the same quadratic is used for any random variable with mean π and second moment in the interval $[\pi^2, 15.915]$. Only the weights for the optimal distribution change, and hence the lower bound, too, changes. (Note: This is obvious since the bound is $-0.2665 + 0.8063\pi - 0.1283\mu_2$.) The weights are p, $1-2p$, and p at 0.6829, π, and 5.6003 where $p = \frac{1}{2}\frac{\mu_2-\pi^2}{(0.6829-\pi)^2}$. (Note: This illustrates that one quadratic may correspond to several optimal distributions. Similarly, a distribution may correspond to a multitude of quadratics.)

Another question arises when $d \leq 0$. Then a quadratic passing through $(0,0)$, (π,d), and $(2\pi,0)$ will be optimal. Compare this with a solution for $\max\limits_{F\in\mathcal{F}} \int_0^{2\pi} (1-\cos(x))dF(x)$. Here an optimal quadratic for any μ_1,μ_2 pair is the one passing through $(0,0)$, $(\pi,2)$, and $(2\pi,0)$ i.e., $-\frac{2}{\pi^2}(x-\pi)^2 + 2$. This is clear since any μ_1,μ_2 pair can be obtained by placing the appropriate weights at $x = 0$, $x = \pi$, and $x = 2\pi$. $\square$

This example was not designed to show every difficulty that could be encountered. Rather, it was intended to show that, when g does not satisfy a global (i.e., for all μ_1,μ_2) feasibility condition (such as g' convex), one must verify feasibility for the particular

μ_1,μ_2 pair at hand. The ideas presented in chapter two and in the present section are efforts to overcome this. They fall short of a coherent system, however, and at present the only general framework that encompasses them both is the formulation of Karlin and Studden (1966), or the less structured point of view in David and Kim (1979), concentrating just on the general property of weak duality, without exploring conditions for existence. Hopefully, an intermediate level exists. Until such a development, however, the two methods mentioned should prove useful for a wide variety of applications.

5. Applications in Utility

In the previous chapter, utility theory was discussed in light of results developed in this work. This section presents further specific applications for utility theory.

The first example examines the basic ideas of the E-V analysis discussed in Chapter one with the aid of some elementary random variables and bounds for expected utility. The second example extends the example of Mantell (1976), including the Levy-Markowitz f_k approximator, and imposes an upper bound on stochastic payoff. The last example is drawn from a recent study of nuclear reactor safety. There it is seen that calculating bounds for expected utility is pertinent to multi-attribute utility analysis when payoffs are known to only the first two moments. This will not be covered in detail, however, preserving the illustrative spirit of this chapter.

Traditional E-V (mean-variance) analysis has strong geometrical motivation. Suppose that the set of all possible (μ,σ^2) pairs constitutes the region shown in Figure 4.5.1.

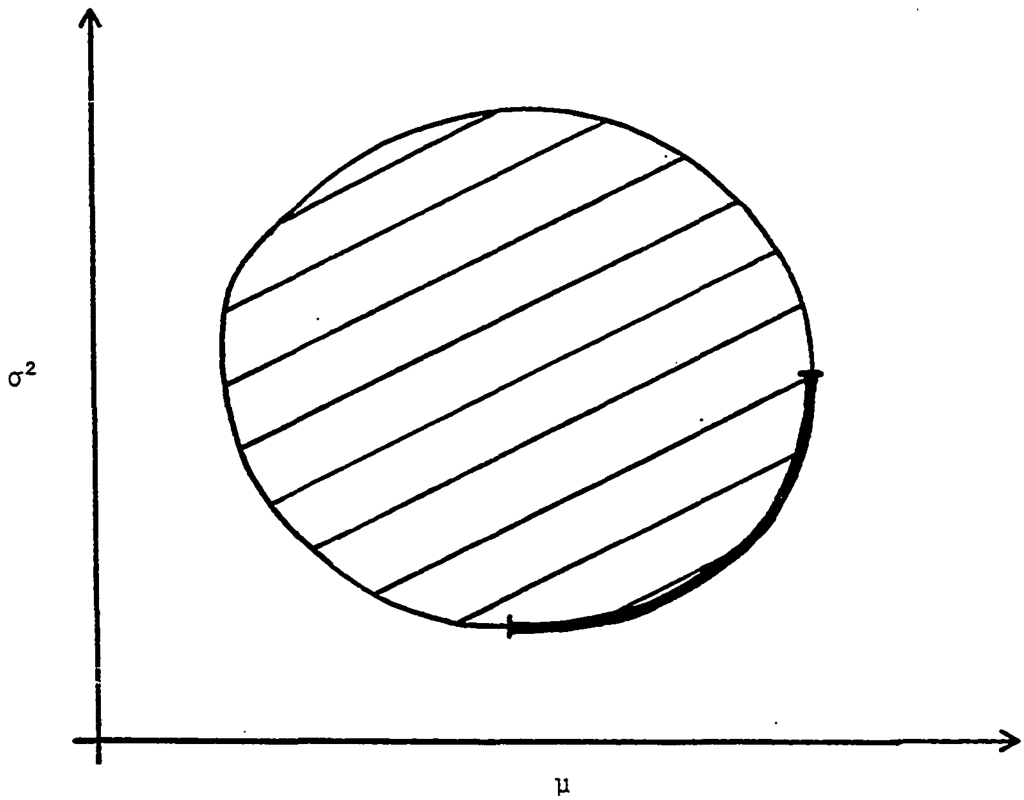


Figure 4.5.1. A possible feasible region for (μ,σ^2)

The lowermost right edge of the region is called the E-V efficient boundary or frontier. An investor would typically desire a payoff distribution whose (μ,σ^2) pair lies on this boundary.

The following notions are often inferred by practitioners of the theory:

1) For a given mean, smaller variance is better.

2) For a given variance, larger mean is better.

3) Increased mean and decreased variance is better.

Use of one of the traditional mean-variance approximators is consistent with these guidelines. One should be aware, however, that when the set of choices is limited, then a (μ,σ^2) point corresponding to the optimizing distribution is not necessarily to be found on the E-V efficient boundary. The following example illustrates this point.

The reader wanting more details is referred to Proposition 1 in Baron (1977).

<u>Example 4.5.1</u> Suppose that an investor has the utility function

$$u(x) = \sqrt{x} \quad \text{on} \quad [0,1] \ .$$

Consider the following five random payoffs.

$$X_1 = \begin{cases} 0.4 & \text{wp} \quad 5/6 \\ 1.0 & \text{wp} \quad 1/6 \end{cases}$$

$$X_2 = \begin{cases} 0 & \text{wp} \quad 1/6 \\ 0.6 & \text{wp} \quad 5/6 \end{cases}$$

$$X_3 = \begin{cases} 0.01 & \text{wp} \quad 1/6 \\ 0.61 & \text{wp} \quad 5/6 \end{cases}$$

$$X_4 = \begin{cases} 0.02 & \text{wp} \quad 1/7 \\ 0.58 & \text{wp} \quad 6/7 \end{cases}$$

and $$X_5 = \begin{cases} 0 & \text{wp} \quad 1/7 \\ 0.6 & \text{wp} \quad 6/7 \end{cases}$$

Table 4.5.1 summarizes the information for these random variables.

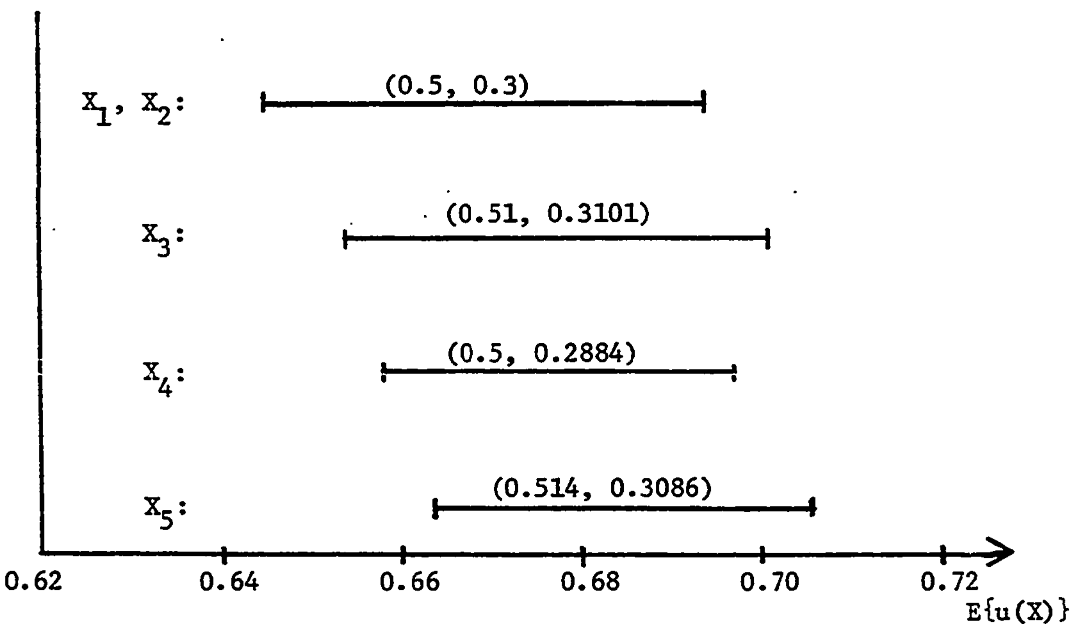
Table 4.5.1. Expectations for random variables

Variable	μ_1	μ_2	σ^2	$E\{u(X)\}$
X_1	0.500	0.3000	0.0500	0.694
X_2	0.500	0.3000	0.0500	0.645
X_3	0.510	0.3101	0.0500	0.668
X_4	0.500	0.2884	0.0384	0.673
X_5	0.514	0.3086	0.0441	0.664

Several features are of interest. Note that $E\{X_3\} > E\{X_1\}$ and $\text{Var}\{X_3\} = \text{Var}\{X_1\}$, yet $E\{u(X_3)\} < E\{u(X_1)\}$, contradicting 2) above. Further, $E\{X_4\} = E\{X_1\}$ and $\text{Var}\{X_4\} < \text{Var}\{X_1\}$ but $E\{u(X_4)\} < E\{u(X_1)\}$, contradicting 1) above. Finally, X_1 and X_5 contradict 3) since $E\{X_5\} > E\{X_1\}$ and $\text{Var}\{X_5\} < \text{Var}\{X_1\}$ yet $E\{u(X_5)\} < E\{u(X_1)\}$.

The point of course is that, depending on the utility function, and the details of the distribution of the random payoff, 1), 2) and 3) may be fallible guidelines when "better" is interpreted as "having greater expected utility".

But how does Theorem 2.6.1 aid the decision maker? For u' convex, as in this example, an interval indicating the possible values for $E\{u(X)\}$ may be plotted as a function of the moments μ_1, μ_2 of the unknown random payoff. As can be seen in Figure 4.5.2, the overlap is quite substantial for the μ_1, μ_2 pairs in the present example. (See Example 4.5.3 for a situation where this technique proves much more useful.) If for no other reason, a plot of this type intrinsically reflects the uncertainty involved in the problem of approximating $E\{u(X)\}$.



(μ_1, μ_2) given above segment.

Figure 4.5.2. Intervals for $E\{u(X)\}$

Another use concerns the situation where some distributions are known and some are unknown. For example, suppose that the distribution of X_2 is completely known while the other X's are known only to μ_1 and μ_2 (as given). From the plot (or simply looking at the lower bounds) it is clear that the expected utility for X_3, X_4, and X_5 must be better, and that for X_1 can be no worse than the expected utility for X_2 . More generally stated, a two point distribution placing mass at the left endpoint of the interval has expected utility less than that of any other distribution with the same mean and variance. In similar fashion, in the example X_1 has the maximum expected utility among all distributions with mean 0.5 and variance 0.05 . □

Section four of chapter three briefly discussed the work of Mantell (1976). In that paper, an example was presented which will be expanded upon in the following example. The Levy-Markowitz f_k approximator will be incorporated in the analysis. The effect of adding an upper bound for the random variable will also be examined.

<u>Example 4.5.2</u> An investor has a fixed amount of money to invest in two mutual funds. Let R_1 and R_2 be the random variables of rate of return for the two funds. Let w be the proportion of the money allocated to the first fund, $0 \leq w \leq 1$. Then

$$X_w = wR_1 + (1-w)R_2$$

is the random variable of interest, being the total rate of return for the investment.

Suppose that the investor has the utility function

$$u(x) = 1-e^{-x}, \quad x \geq 0 . \tag{4.5.1}$$

Suppose also that

$$E\{R_1\} = 0.05$$

$$E\{R_2\} = 0.08$$

$$\mathrm{Var}\{R_1\} = 0.0225$$

$$\mathrm{Var}\{R_2\} = 0.16$$

and $\mathrm{Cov}(R_1,R_2) = 0.03$.

This implies that

$$\begin{aligned} \mu_w = E\{X_w\} &= 0.05w + 0.08(1-w) \\ &= 0.08 - 0.03w \end{aligned} \tag{4.5.2}$$

and

$$\begin{aligned} \sigma_w^2 = \mathrm{Var}\{X_w\} &= 0.0225w^2 + 0.16(1-w)^2 + 0.06w(1-w) \\ &= 0.1225w^2 - 0.26w + 0.16 . \end{aligned} \tag{4.5.3}$$

Theorem 2.6.1 applies, so that the lower bound for $E\{u(X_w)\}$ may be written as

$$\ell(w) = \frac{\left\{1 - \exp\left(-\frac{\sigma_w^2 + \mu_w^2}{\mu_w}\right)\right\}(\mu_w^2)}{\sigma_w^2 + \mu_w^2} . \tag{4.5.4}$$

This function is plotted in Figure 4.5.3. The Levy-Markowitz f_k approximator, discussed in section five of chapter three, is also

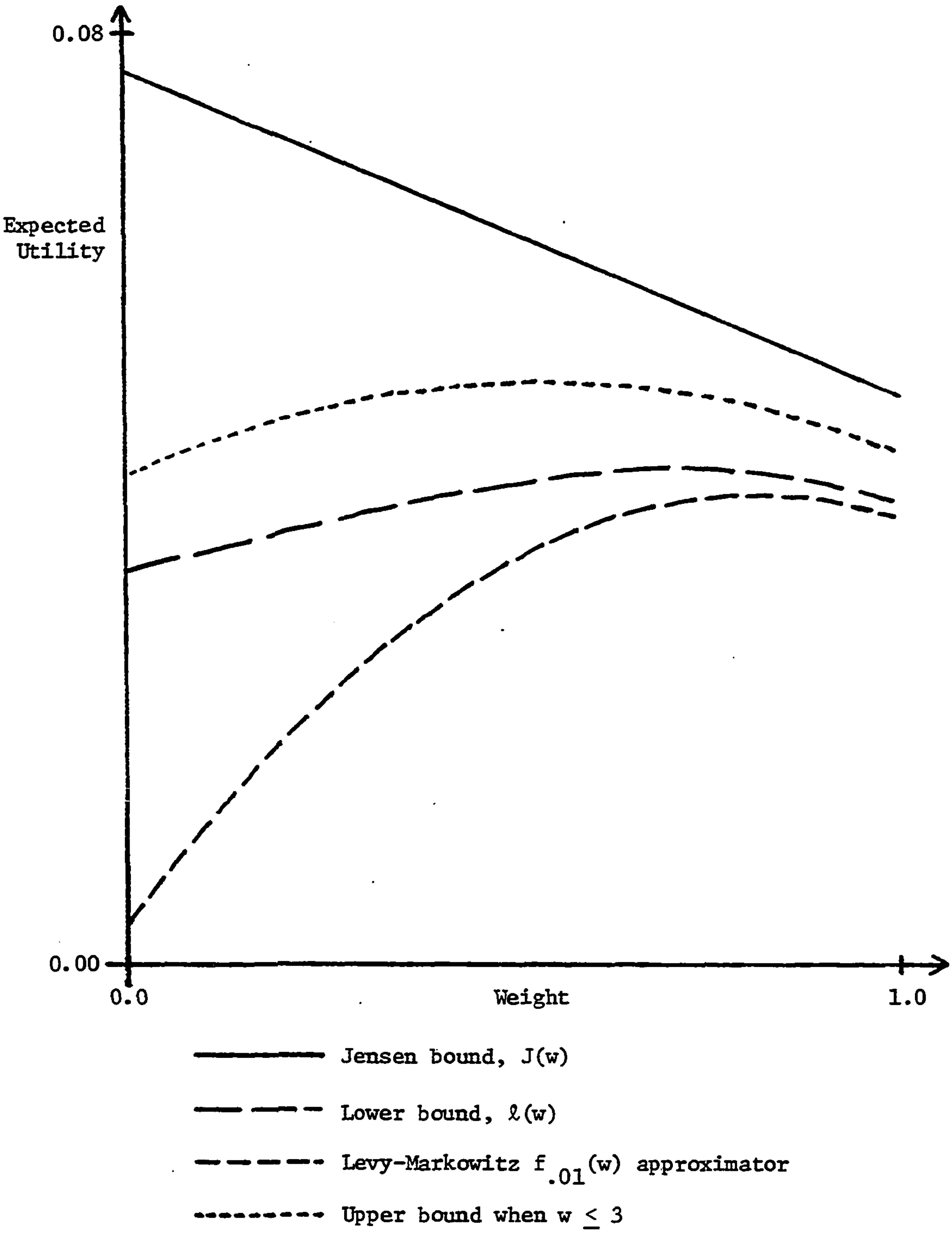


Figure 4.5.3. Extension of Mantell example

plotted. As can be seen, that approximator lies below $\ell(w)$ on the entire interval, and markedly so for small w. Furthermore, note that X cannot be a two point distribution unless $w = 1$, $w = 0$, or one of the two rates of return has a degenerate distribution. The bound is attainable only if X has a two point distribution.

Two upper bounds are also plotted. The first is the Jensen bound, given by

$$J(w) = 1 - \exp(-\mu_w) . \tag{4.5.5}$$

The other is obtained from Theorem 2.6.1 by arbitrarily imposing an upper bound of 3 on X. This corresponds to an assumption that the investor will do no better than a 400% rate of return on the investment. Such an assumption should not be made lightly. The investor should question this kind of assumption carefully since its failure invalidates the bound. Imposing any finite bound on X will yield an improvement over the Jensen bound, of course, though it may be only a slight advantage if taken unreasonably large.

Before leaving this example, it seems worthwhile to briefly mention the more general case when

$$X = \sum_{i=1}^{n} w_i R_i ,$$

$$\sum_{i=1}^{n} w_i = 1 , \quad w_i \in [0,1] \; \forall \, i .$$

As Mantell pointed out, finding the set of w_i's yielding the maximum lower bound generally entails solving a set of n-1 non-linear equations for n-1 variables. This may or may not be difficult. Even if difficult to solve exactly, it is relatively inexpensive to generate lower bound values for a grid of w_i values as Mantell did.

The lower bound function is generally well-behaved so this can be expected to work well. Additionally, the permissible values for the w_i's may be restricted. For example, perhaps some bonds under consideration are available only in one hundred dollar increments. Note also for the case of n=3 that a contour plot of the lower bound may easily be produced. $\square$

The last example of this section (indeed, of this thesis) concerns an actual application of utility theory. It seems particularly fitting to end this thesis with a brief view of some of the work which inspired it.

<u>Example 4.5.3</u> Nuclear Reactor Safety. Multiattribute utility theory was used to evaluate seven research and development programs for nuclear reactor safety in Ritzman and Husseiny (1980). The computer package used for the analysis, MAUP, employed a derivative technique to estimate the uncertainties involved in the expected utilities (see Hale (1980)). This actually amounted to little more than the usual E-V approximation. The present example presents one of the subattributes considered and applies Theorem 2.6.1 to obtain bounds for expected utility.

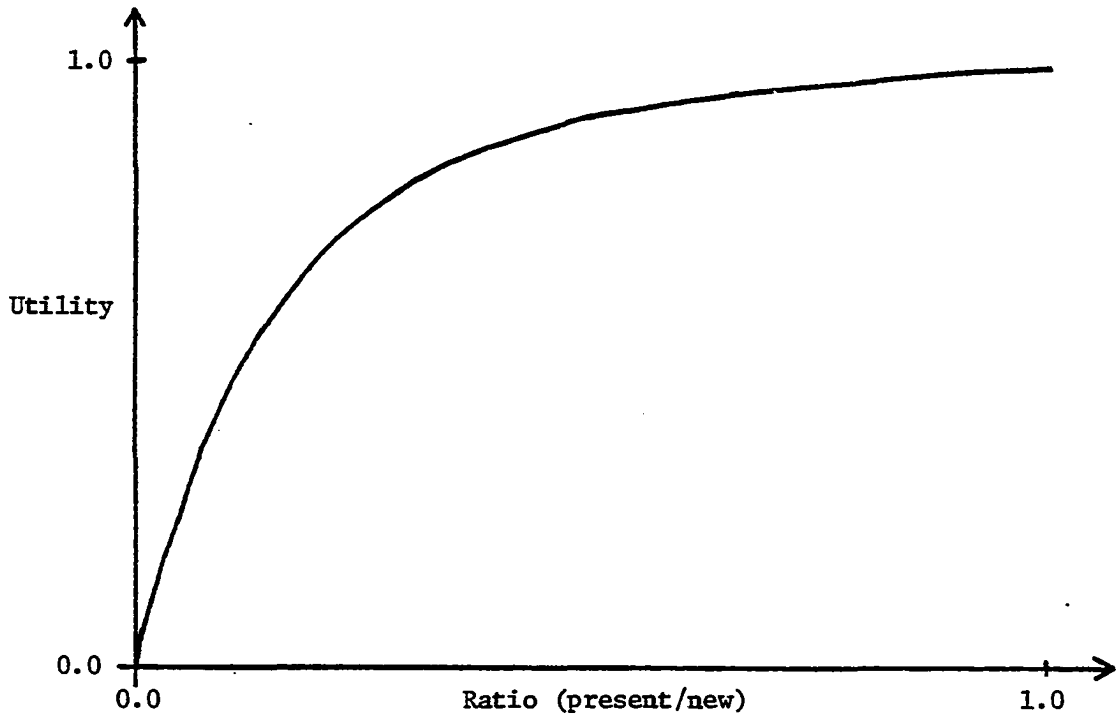


Figure 4.5.4. Utility curve for risk reduction potential

A major area of concern is risk reduction potential. One of the aspects of this concerns the failure of operations personnel to perform as necessary to prevent accidents. This was measured as the ratio of present expected human error rate to the anticipated expected human error rate under a new R&D program. The best and worst values were taken to be 10 and 0.5 respectively. All values were scaled to [0,1] by

$$x = \frac{\text{ratio} - 0.5}{9.5}$$

It was found that the utility function

$$u(x) = \frac{1.40 - \exp(-6.96x) - 0.40 \exp(-0.75x)}{1.21} \tag{4.5.6}$$

described the evaluator's preference structure quite well. See Figure 4.5.4. The derivative is convex so that

$$\frac{\mu_1^2}{\mu_2} u(\mu_2/\mu_1) \leq E\{u(X)\} \leq \frac{\mu_2 - \mu_1^2 + (1-\mu_1)^2 u\left(\frac{\mu_2-\mu_1}{\mu_1-1}\right)}{(\mu_2-\mu_1^2) + (\mu_1-1)^2} \tag{4.5.7}$$

The information for the seven programs is summarized in Table 4.5.2.

Table 4.5.2. Utility bounds for the seven programs

Program	μ_1	μ_2	σ^2	lower bound	upper bound
1	0.259	0.08642	0.01920	0.494	0.517
2	0.037	0.00412	0.00274	0.092	0.097
3	0.015	0.00066	0.00044	0.040	0.041
4	0.081	0.01663	0.00999	0.181	0.195
5	0.026	0.00119	0.00052	0.070	0.071
6	0	0	0	0	0
7	0.004	0.00004	1.43×10^{-8}	0.011	0.011

A plot such as Figure 4.5.2 can easily be produced, yet a simple visual inspection of the table reveals that the intervals are all well separated. The programs may now be ranked with a high degree of confidence. Traditional methods would come to the same conclusions in this example, but one clearly would not have the same degree of confidence in one's conclusions.

Perhaps more important are the implications for multiattribute utility. An R&D program is not selected on one attribute alone. The present attribute is only one of the fifty-two used in the study. Several of these had a convex decreasing utility function. The traditional methods overestimated expected utility for the curve presented here by quite a large amount. For example, $f_{.01} = .747$ for program 1. Unfortunately, the traditional methods will underestimate for most convex decreasing utility functions. The effect of combining these overestimates and underestimates in an additive or multiplicative multiattribute utility function is unclear at the present time. It could certainly give rise to misleading results.

A safer course would probably be to take the midpoint of each interval as an estimate. Note also that a worst case analysis is also quite simple now by utilizing the lower bounds. Admittedly it might be quite unrealistic to set all values at their best or worst, but it could provide conservative bounds for the overall expected utility of each program. This could, of course, prove useful in evaluating the performance to expect from a given program and certainly seems worthy of future research. $\square$

6. Conclusions

As pointed out in section four, this work is a point of beginning, rather than of ending. The emphasis has been on the case where one has a finite endpoint for the region of interest and an optimal distribution placing mass at that endpoint. More work needs to be done considering what happens when this is not the case. Continuous and unimodal distributions are enticing assumptions for investigation.

Other possibilities include more fully investigating the implications of this work. One area, mentioned in the previous section and having great promise, is multiattribute utility theory. The need for this is rather compelling, since uncertainties present there inspired the present work. Another field is reliability, as pointed out in section two.

Rather than reiterate the detailed conclusions and suggestions found throughout this work it seems a better course to end with some elements of the general point of view implicit in it. Any answer should be tempered by an awareness of assumptions. If constraints are known, they should be used to advantage. Force any analysis to a reasonable neighborhood of reality, and succeeding analyses to even smaller such neighborhoods.

V. REFERENCES

Amihud, Y. 1979. Critical examination of the new foundation of utility. Pages 149-160 in M. Allais and O. Hagen, eds. Expected Utility and the Allais Paradox. D. Reidel Publishing Company, Dordrecht, Holland.

Arrow, K. J. 1974. Use of unbounded utility functions in expected-utility maximization-response. Quarterly Journal of Economics 88:136-138.

Baron, D. P. 1977. On the utility theoretic foundations of mean-variance analysis. Journal of Finance 32:1683-1697.

Bawa, V. S., E. B. Lindenberg, and L. C. Rafsky. 1979. An efficient algorithm to determine stochastic dominance admissible sets. Mgmt. Sci. 25:

Becker, G. M., M. H. DeGroot, and J. Marschak. 1964. Measuring utility by a single-response sequential method. Behavioral Science 9:226-232.

Berger, R. W. and M. D. Hale. 1980. Using programmable calculators in quality control -- tutorial. 34th Annual ASQC Technical Conference Transactions, Chelsea, Michigan.

Bernoulli, D. 1954. Exposition of a new theory of the measurement of risk. Econometrica 22:23-36. (Translation of a paper originally published in Latin in St. Petersburg in 1738.)

Biorn, E. 1974. Estimating flexibility of marginal utility of money--errors-in-variables approach. Eur. Econ. R. 5:177-185.

Borch, K. 1979. Utility and stochastic dominance. Pages 193-201 in M. Allais and O. Hagen, eds. Expected utility and the Allais Paradox. D. Reidel Publishing Company, Dordrecht, Holland.

Brumelle, S. L. and R. G. Vickson. 1975. A unified approach to stochastic dominance. Pages 101-113 in W. T. Ziemba and R. G. Vickson, eds. Stochastic Optimization Models in Finance. Academic Press, New York.

Chipman, J. S. 1973. The ordering of portfolios in terms of mean and variance. Rev. Econ. Studies 40:167-190.

David, H. T. and G. H. Kim. 1979. Pragmatic optimization of information functionals. Pages 167-181 in J. S. Rustagi. Optimizing Methods in Statistics. Academic Press, New York.

Davidson, D., P. Suppes, and S. Siegel. 1957. Decision Making: An Experimental Approach. Stanford University Press, Stanford, Cal.

DeFinetti, B. 1979. A short confirmation of my standpoint. Page 161 in M. Allais and O. Hagen, eds. Expected utility and the Allais Paradox. D. Reidel Publishing Company, Dordrecht, Holland.

Eggleston, H. G. 1958. Convexity. Cambridge tracts in mathematics and mathematical physics 47. Cambridge University Press, Cambridge.

Fishburn, P. C. 1964. Decision and Value Theory. J. Wiley & Sons, New York.

Fishburn, P. C. 1968. Utility theory. Mgmt. Sci. 14:335-378.

Fishburn, P. C. 1969. A general theory of subjective probabilities and expected utilities. Ann. Math. Stat. 40:1419-1429.

Fishburn, P. C. 1976. Theory versus practice in risk analysis: An empirical study: A comment. The Accounting Review 51:657-662.

Fishburn, P. C. and R. G. Vickson. 1978. Theoretical foundations of stochastic dominance. Pages 39-113 in G. A. Whitmore and M. C. Findlay, eds. Stochastic Dominance. D. C. Heath and Company, Lexington, Mass.

Greer, W. R. Jr. 1976. Theory versus practice in risk analysis: An empirical study: A reply. The Accounting Review 51:663.

Hadar, J. and W. R. Russell. 1969. Rules for ordering uncertain prospects. Amer. Econ. Rev. 49:25-34.

Hagen, O. 1979. Towards a positive theory of preferences under risk. Pages 271-302 in M. Allais and O. Hagen, eds. Expected utility and the Allais Paradox. D. Reidel Publishing Company, Dordrecht, Holland.

Hale, M. D. 1980. User's manual for the multi-attribute utility package (MAUP). SAI Technical Report SAI-101-81-AM. (Science Applications, Inc., Ames, Iowa).

Isii, K. 1960. The extrema of probability determined by generalized moments (I) bounded random variables. Ann. Inst. Stat. Math. 12:119-133.

Isii, K. 1963. On sharpness of Tchebycheff-type inequalities. Ann. Inst. Stat. Math. 14:185-197.

Isii, K. 1964. Inequalities of the types of Chebyshev and Cramér-Rao and mathematical programming. Ann. Inst. Stat. Math. 16:277-293.

Karlin, S. and W. J. Studden. 1966. Tchebycheff Systems: With Application in Analysis and Statistics. J. Wiley & Sons, New York.

Keeney, R. L. and H. Raiffa. 1976. Decisions with multiple objectives: Preferences and value tradeoffs. J. Wiley & Sons, New York.

Kim, G. H. 1978. Constructive optimization of infinite dimensional problems in probability and statistics. Ph.D. Thesis. Iowa State University, Ames, Iowa.

Kim, G. H. 1979. Bounds for strength-stress interference via mathematical programming. To be published in Nav. Log. Quart.

Kim, G. H. and H. T. David. 1979. Large deviations of functions of Markovian transitions and mathematical programming duality. Ann. Probability 7:874-881.

Kingman, J. F. C. 1963. On inequalities of the Tchebychev type. Proc. Camb. Phil. Soc. 59:135-146.

Leung, P. 1978. Sensitivity analysis of effect of variations in form and parameters of a multi-attribute utility model-survey. Behavioral Science 23:478-485.

Levy, H. and H. M. Markowitz. 1979. Approximating expected utility by a function of mean and variance. Amer. Econ. Rev. 69:308-317.

Loistl, O. 1976. The erroneous approximation of expected utility by means of a Taylor's series expansion: Analytic and computational results. Amer. Econ. Rev. 66:904-910.

MacCrimmon, K. R. and S. Larsson. 1979. Utility theory: Axioms versus 'paradoxes'. Pages 333-409 in Expected Utility and the Allais Paradox. M. Allais and O. Hagen, eds. D. Reidel Publishing Company, Dordrecht, Holland.

Mallows, C. L. 1956. Generalizations of Tchebycheff's inequalities. J. Roy. Stat. Soc., Ser. B. 18:139-176.

Mantell, E. H. 1976. The applicability of the maxi-min principle to portfolio selection in terms of means and variances of unknown distributions. Dec. Sci. 7:175-185.

Markowitz, H. M. 1952. Portfolio selection. Journal of Finance 7:77-91.

Marshall, A. W. and I. Olkin. 1960a. A one-sided inequality of the Chebyshev type. Ann. Math. Stat. 31:488-491.

Marshall, A. W. and I. Olkin. 1960b. Multivariate Chebyshev inequalities. Ann. Math. Stat. 31:1001-1014.

Mosteller, F. and P. Nogee. 1951. An experimental measurement of utility. J. Polit. Econ. 59:371-404.

Porter, R. B., J. R. Wait, and D. L. Ferguson. 1973. Efficient algorithms for conducting stochastic dominance tests on large number of portfolios. J. Finan. Quant. Anal. 8:71-78.

Pukelsheim, F. 1978. A quick introduction to mathematical programming with applications to most powerful tests, non-negative variance estimate and optimal design theory. Technical Report No. 128. (Department of Statistics, Stanford University, Stanford, California).

Ramsey, F. P. 1931. Truth and probability. Pages 156-198 in R. Braithwaite, ed. The oundations of Mathematics and Other Logical Essays. Routledge and Kegan Paul, London.

Ritzman, R. L. and A. A. Husseiny. 1980. A selection methodology for LWR safety programs and proposals - Volume 2 - Methodology application. SAI Technical Report SAI-195-80-PA. (Science Applications, Inc., Ames, Iowa).

Rockafellar, R. T. 1974. Conjugate Duality and Optimization. SIAM, Philadelphia.

Rogosinski, W. W. 1958. Moments of nonnegative mass. Proc. Roy. Soc., Ser. A. 245:1-27.

Russell, W. R. and Tae Kun Sev. 1978. Admissible sets of utility functions in expected utility maximization. Econometrica 46:181-184.

Samuelson, P. A. 1970. Mean-variance and safety first approaches and their extensions. Rev. Econ. S. 37:537-542.

Sposito, V. 1975. Linear and Nonlinear Programming. The Iowa State University Press, Ames, Iowa.

Tesfatsion, L. 1976. Stochastic dominance and maximization of expected utility. Rev. Econ. Studies 43:301-315.

von Neumann, J. and O. Morgenstern. 1947. Theory of Games and Economic Behavior. Princeton University Press, Princeton, N.J.

Whitmore, G. A. 1970. Third-degree stochastic dominance. Amer. Econ. Rev. 60:457-459.

VI. ACKNOWLEDGEMENTS

I want to thank the people who helped throughout this work. I am grateful to Darlene Wicks, who showed patience throughout the typing, and especially Herbert T. David whose hard work, guidance, and editorial assistance were invaluable. The many faculty members at Iowa State University who stimulated and instructed me are also greatly appreciated, as well as my college instructor, Robert McMillan, who reinforced my desire to pursue an advanced degree. I also thank my supervisor, Cliff Rudy, for his patience and understanding. Lastly, I thank my wife, Vernona, and my children, Crystal and Jennifer, for their encouragement and love. They provided inspiration when the writing was difficult and my spirits low.

This work was partially supported by AFOSR Grant #78-3518.